

Journal of Language Teaching and Research

ISSN 1798-4769 (print); ISSN 2053-0684 (online)

Volume 17, Number 3, May 2026

Contents

REGULAR PAPERS

- Describing Bilinguals Through a Conceptual Bridge: Integrating Second Language Acquisition Into Bilingualism Research 757
Mackarena Kartsevski
- Why Are Students Failing Listening? Investigating Reasons for the Low Performance of Omani General Foundation Program Students in English Listening Exams 767
Fakhriya H. Al-Rashdi
- Bilingual Texting in the Digital Age: Arabic–English Code-Switching in SMS 778
Wasan Khalid Ahmed, Ibtihal Hassan Mussa, Laith Salman Hadla, Marwan Harb Alqaryouti, and Ala Eddin Sadeq
- Documenting Linguistic Diversity: An Analysis of Compound Word Formation in the Goris Armenian Dialect 789
Lusine R. Nersisyan, Mher S. Kumunts, Ani A. Ghazaryan, Manuk H. Falakyan, Armine A. Mkrtumyan, Leonid G. Telyan, Anahit S. Galstyan, Susanna K. Tioyan, Harout K. Marachlian, Tereza S. Shahverdyan, Anna H. Abajyan, Manana G. Dalalyan, Vanuhi M. Baghramyanyan, and Karine L. Arakelyan
- Modelling Cognitive and Metacognitive Strategies and Their Relationships to Reading Comprehension Test Performance of Thai High School EFL Learners 801
Panassanan Kitichaidateanan, Nithipong Yothachai, and Apisak Sukying
- "I Am in Control": Strategic Gender Performance and Agency in Joyce Carol Oates's *Marya: A Life* 812
Lingling Xu, Manimangai Mani, and Hardev Kaur a/p Jujar Singh
- Fear-Driven Dialogue: A Linguistic Study of *Death Knocks* and *Death* 821
Mohammed A. Abou Adel, Ayman I. Elhalafawy, Samir H. Khalifa, Ahmad M. Al Mahamed, and Fatma T. El-Zaghal
- Revisiting Border and Defying Boundaries: Unravelling Echoes of the Partitioned Woman in Geetanjali Shree's *Tomb of Sand* 834
Bharti Sharma and Sarika Tyagi
- Rooted Kinship and Storied Matter: Relational Ontologies and the Ethics of Coexistence in Elif Shafak's *The Island of Missing Trees* 842
Farhan Ahmad, Shamsudheen Mannenkuzhiyan, Wahaj Unnisa Warda, and Shehnoor Shan
- Language Change in Social Media and Its Impact on the Linguistic Identity of Indonesia's Younger Generation: A Sociolinguistic Study 850
M. Ridha Ajam, Mujahid Taha, and Andi Sumar-Karman
- The Bird Image in English and Arabic Poetry: A Stylistic Comparative Study of Selected Poems 862
Amjad Mohammad Almuqheed, Asma Hussein, Emad Abdallah Abuhammam, and Ali Bani Ali
-

Appraisal Framework Instruction and Its Impact on English Writing: A Study Among Vietnamese University Students <i>Hong B. Nguyen</i>	869
The Impact of the Flipped Classroom Model on Developing EFL Students' Four Language Skills: An Experimental Study <i>Rasha Dorgham and Lamya Bin Obiad</i>	882
Media Discourses of Moderation: A Critical Discourse Analysis of Global Coverage of Ahmed Al-Sharaa's Presidency <i>Ibrahim Bani Abdo, Abdelhameed Al Awabdeh, Mohammad Albtoush, Abeer Hameed Albashtaw, Nabil Al-Awawdeh, and Mheel Al-Smaihyeen</i>	893
Leveraging Artificial Intelligence to Support Language Processing in Children With Dyslexia and Dysgraphia: A Design-Based Study <i>Sri Sukasih, Atip Nurharini, and Andarini Permata Cahyaningtyas</i>	902
The Role of Tone in Dubbing Children's Cartoons: A Case Study of <i>Timon & Pumbaa</i> Series <i>Alshaymaa Y. Alharbi</i>	915
Exploring Students' Perspectives on Quizlet as a Gamified Vocabulary Learning Tool in a Business English Course <i>Thuan D. Pham and Tam T. Pham</i>	925
Ecocentrism, Irony, and the Wounds of Nature in the Poem "Aku Ini": Ecocriticism in the Context of North Maluku <i>Wildan, Ramis Rauf, Ratna, Afriani Ulya, Vidhia Agmareina Hirto, Sulmi Magfirah, Fitriarningsih Pratiwi Mahmud, Dhini Yustia Widhyah Saputri, Ulfah Dwi Januarty, and Ahmad Rifani Talaohu</i>	935
Exploring Lexical Processing and Sustainability Awareness in EFL Learners: A Psycholinguistic Discourse-Cognition Perspective <i>Mohammad Abdulaal, Ahmad Sharaf Eldin, Hanan Al-Johani, and Mohammad Aldawoody</i>	945
Video-Based Platforms and Speaking Skill Development in Undergraduate EFL Contexts: Insights From Flipgrid <i>Maytham Hashim Mutlaq, Harmi Izzuan bin Baharum, Ali Hamed Barghi, and Seriaznita Mat Said</i>	955
The Effect of Learning Modalities and Self-Regulated Learning on EFL Learners' Critical Thinking Development <i>Rosmania Rima, Yudi Juniardi, and Syafrizal Syafrizal</i>	965
Saudi University EFL Students' Understanding of Plagiarism: Perceived Knowledge, Recognition Patterns, and Conceptual Ambiguity in Academic Writing <i>Nawal I. Alhodithi</i>	978
The Impact of AI-Assisted Tools on Enhancing the Quality of Legal Translation Outputs: A Case Study of Jordanian Students <i>Alalddin A. Al-Tarawneh, Mohammed Al-Badawi, Muath Algazo, Rula Ahmad Mahmoud Abu-Elrob, Bilal Alsharif, and Nader Albkower</i>	988
Enhancing English Speaking Skills of Vietnamese University Students Through ChatGPT-Aided Mind Maps <i>Thi Huong Thao Dang</i>	999
EFL Policy Alignment With Saudi Arabia's Economic Transformation and Vision 2030: A Systematic Review <i>Saad Eid Albaqami, Dhaif Alzahrani, Mohammad Jamshed, and Mohammad Rezaul Karim</i>	1008

A DQF-MQM Evaluation of Machine Translation in English–Arabic Customs Law <i>Karam Damseh and Mozghan Ghassemiazghandi</i>	1019
Authorial Ideology and Worldview in Kijne's <i>Kota Emas</i> : A Genetic Structuralist Analysis <i>Insum Malawat, Hengki, and Burhanuddin Arafah</i>	1028
Symbolic Articulations of Post-Traumatic Memory in Women's Poetry: Fire, Ashes, and Self-Transformation in Atwood and Al-Malaika <i>Fatimah A. Alwadaani</i>	1039
Attitudinal Resources in the English Textbook <i>Life – Upper Intermediate</i> : A Systemic Functional and Intercultural Perspective <i>Ngoc T. B. Han</i>	1048
The Impact of Cultural Nuances in the Translation of Indonesian Government Social Media Content on the Digital Literacy of Young Generations <i>Burhanuddin Arafah, Herawaty Abbas, Muhammad Hasyim, Muhammad Ridha Anugrah Latief, and Taqdir</i>	1057
Gendered Variations in Translanguaging Practices and Attitudes by EFL Teachers in Saudi Arabia <i>Sita N. Aldossari and Manar Almanea</i>	1070
Sisterhood in Leila Aboulela's <i>Bird Summons</i> and Toni Morrison's <i>Sula</i> <i>Nadera Nader Al-Staity, Nasaybah Walid Awajan, and Sahar Mohammad Tarawneh</i>	1081
Figurative Language and Aesthetic Value in the Dance Drama <i>The Search for Sita</i> <i>Ni Ketut Dewi Yulianti, Ni Nyoman Kasih, I Wayan Suardana, Ni Putu Tisna Andayani, Sang Ayu Isnu Maharani, Ni Made Ayu Widiastuti, and Agus Darma Yoga Pratama</i>	1090
Language and Digital Content in Saudi E-Commerce: An Analytical Review <i>Mohammad Abdulaal and Mohammad Aldawoody</i>	1100
From Brainstorming to Boundaries: Writing Center Tutors' Use of Generative AI <i>Maria Eleftheriou and Muhammad Ahmer</i>	1110
Voices on the Margins: An Intersectional Study of Women Oppression and Agency in Jasmin Darznik's Memoir <i>The Good Daughter</i> <i>Aishwarya Mary Thomas and Mohamed Sahul Hameed M. A</i>	1122

Describing Bilinguals Through a Conceptual Bridge: Integrating Second Language Acquisition Into Bilingualism Research

Mackarena Kartsevski

Faculty of Educational Sciences, Universidad de Talca, Linares, Chile;
Millennium Nucleus for the Science of Learning (MiNSoL), Chile

Abstract—Over the last decade, research on bilingualism has placed increasing importance on the accurate and detailed description of bilingual experiences. This shift has arisen as bilingualism is now understood as a continuum rather than a dichotomous condition, emphasizing the dynamic experiences that shape the bilingual mind. While the field has made substantial progress in describing linguistic and contextual variables, it still lacks a coherent framework to capture the full range of individual differences. This paper adopts a conceptual synthesis approach (Jaakkola, 2020) to build a bridge between Bilingualism and Second Language Acquisition (SLA)—two related, but traditionally separate fields, that share a common object of study: individuals who learn and use more than one language. Drawing on SLA's established frameworks for conceptualizing and measuring individual differences—such as proficiency, motivation, aptitude, and personality—this synthesis explores how these constructs can enrich the ecological and experience-based understanding of bilingualism. Furthermore, this work situates this integration within the broader framework of the Science of Learning, a transdisciplinary field that connects cognitive science, linguistics, psychology, and education to explain learning as an adaptive, context-sensitive process. By aligning SLA, bilingualism, and the Science of Learning, the paper advances a view of language learning as a lifelong, dynamic phenomenon that reflects how humans learn, adapt, and transform through experience.

Index Terms—bilingualism, Second Language Acquisition, Science of Learning, individual differences, education

I. INTRODUCTION

Bilinguals are individuals who use two or more languages daily (Bialystok, 2017; de Bruin, 2019; López et al., 2021). As this definition is very broad, it describes a wide variety of people worldwide who are very different from one another. For example, some bilinguals learn two languages from birth, others learn a second language at school, and others learn a community language that is different from the one used at home. These are just a few ways in which people can be bilingual; of course, there are many others.

The importance of these divergent features is that each one shapes how our cognitive system adapts to the experience of learning a second language (Kroll & Rossi, 2023; Pliatsikas, 2023). For this reason, the field of bilingualism has made a tremendous effort to describe them fully, especially because these differences have been thought to explain the distinct results obtained in studies that looked for the bilingual cognitive advantage during the first decade of the 2000s (de Bruin, 2019). One of the many critiques of this alleged advantage has been the poor description of bilinguals, as studies have grouped people under the category of monolingual or bilingual as if they were dichotomous variables (van den Noort et al., 2019).

In this context, the field of bilingualism has moved away from traditional ways of describing bilinguals. In the early days of bilingualism research, individuals who used two languages were classified according to the stage in which they acquired the second language (L2), dividing them into either "early" (during the 'Critical Period', see Lenneberg, 1967 for more details) or "late" (after the 'Critical Period') bilinguals (Paradis & Genesee, 1996). A second traditional classification was based on the order of language acquisition, where bilinguals who acquired the two languages simultaneously were characterized as "simultaneous bilinguals," and those who acquired the L2 after the first was already developed were labeled as "sequential bilinguals" (Swain & Cummins, 1979). A third traditional classification was based on the level of proficiency in each language, distinguishing those individuals with equal proficiency in both languages, labeled as "balanced," from those whose proficiency in one language was more advanced than the other, called "unbalanced" bilinguals (Peal & Lambert, 1962).

Instead of labeling bilinguals and classifying them into categories, research has shown that these descriptions fall short, mainly because we now understand that being bilingual is not a dichotomous variable, but a continuum (Luk & Bialystok, 2013). For these reasons, the field has pointed out that linguistic individual differences, such as age of acquisition, proficiency, language use, and code-switching should be considered to describe a specific sample (Bialystok & Craik, 2021; Bialystok, 2017; de Bruin, 2019; López et al., 2021). Recent developmental research further demonstrates that even

among bilinguals with similar exposure patterns, there is substantial intra-individual variability in linguistic outcomes, highlighting the need for more fine-grained and dynamic descriptions of bilingual experience (Kroll & Rossi, 2023; Meir, 2023; Paradis, 2023).

More recently, environmental variables have gained more attention (e.g., Bratlie et al., 2025; Yow & Li, 2024), especially after the publication of the study by Beatty-Martínez et al. (2020). The authors compared three groups of second language learners in three different contexts. In the first group, bilinguals used their languages in separate contexts and lived in Spain. In the second group, bilinguals used their languages in an integrated context and lived in Puerto Rico. In the third group, bilinguals used their languages in a varied context where they could code-switch if they had the chance. These bilinguals lived in the United States. The results showed that bilinguals recruited cognitive control to regulate their languages differently according to the context in which they operated. While bilinguals in Spain recruited cognitive control reactively, individuals in the U.S. recruited it proactively. Bilinguals in Puerto Rico used a mixture of both. These results shed light on how bilinguals, who, in other studies, could have been grouped together under the label of second language learners, are different from one another.

In this context, the field of bilingualism has made significant progress in describing bilinguals accurately and is now open to incorporating variables that are not strictly linguistic. However, despite this progress, the field still lacks a coherent framework of individual differences that can guide how bilingual experiences are measured and compared across studies. Interestingly, the field of Second Language Acquisition (SLA) has long addressed these same individual differences—such as proficiency, motivation, aptitude, personality, and learning context—and has developed validated instruments and theoretical models to describe them. Yet, these constructs have rarely been systematically adopted in bilingualism research.

Therefore, this paper seeks to build a conceptual bridge between SLA and bilingualism by identifying how SLA-based individual difference constructs can inform the evolving, ecological understanding of bilingualism. The aim is not to conduct an exhaustive review, but to propose a conceptual framework that allows the field of bilingualism to integrate these constructs and move toward more precise and experience-sensitive descriptions of bilingual populations. To accomplish this goal, the study adopts a conceptual synthesis approach (Jaakkola, 2020), described in the following section.

Beyond this disciplinary dialogue, the present work also aligns with the broader principles of the Science of Learning, an emerging field that integrates insights from cognitive science, linguistics, psychology, and education to understand learning as a dynamic, adaptive, and context-sensitive process (OECD, 2018; Bowen et al., 2025; Dumont et al., 2010; Tokuhamma-Espinosa, 2011). From this perspective, both SLA and bilingualism can be seen as complementary lenses for studying how humans learn and adapt through language across the lifespan. The integration of constructs from SLA into bilingualism research—and vice versa—represents a concrete way of advancing this transdisciplinary vision, where learning is understood not only as the acquisition of knowledge, but as a complex interaction among cognition, emotion, and environment (Immordino-Yang & Damasio, 2007; Fischer et al., 2010).

Recent contributions within this field emphasize that advancing the Science of Learning requires connecting theoretical models to contextually grounded practices (Meltzoff et al., 2009; Ferreira & Rodríguez, 2022). In this sense, integrating frameworks from Second Language Acquisition (SLA) and bilingualism offers a concrete example of how transdisciplinary dialogue can enrich both theory and practice. It positions language learning not merely as an instructional process, but as a lifelong, adaptive phenomenon that reflects how humans learn and transform through experience (Ferreira & Rodríguez, 2022; Thomas & Ansari, 2020).

By establishing a conceptual dialogue between bilingualism, SLA, and the Science of Learning, this paper contributes to the ongoing shift toward a more ecological, integrative, and human-centered understanding of bilingual experience. Integrating SLA-derived constructs and measurement tools offers the field a more systematic way to operationalize bilingual experience, while situating this effort within a larger framework that connects language learning to the general principles of how humans learn.

II. METHODOLOGICAL ORIENTATION

This paper follows a conceptual synthesis approach (Jaakkola, 2020) to connect frameworks and constructs across disciplines. Rather than testing empirical hypotheses, the purpose is to articulate how theoretical models and measurement tools from SLA can inform current efforts to describe bilingual experience. This synthesis draws on established constructs within Second Language Acquisition (SLA) that have demonstrated strong theoretical and empirical foundations. The next subsections present these constructs—proficiency, motivation, context of acquisition, aptitude, and personality—from an SLA perspective. By revisiting them conceptually, the paper highlights their potential relevance for advancing the description of bilingual experience. To provide an overview of the constructs that guide this synthesis, Table 1 summarizes the key individual differences described in SLA, their theoretical bases, measurement instruments, and their potential contribution to the study of bilingualism.

TABLE 1
OVERVIEW OF INDIVIDUAL DIFFERENCE CONSTRUCTS IN SLA AND THEIR POTENTIAL RELEVANCE FOR BILINGUALISM

Construct	Definition (SLA perspective)	Key Theoretical Source / References	Measurement Instruments	Potential Relevance for Bilingualism
Proficiency	Communicative ability across skills (reading, writing, listening, speaking) as defined by CEFR.	Council of Europe (2020); Cambridge English (2024)	CEFR, TOEFL, Cambridge Exams	Provides standardized way to describe bilingual competence.
Motivation	Degree of autonomy, competence, and relatedness driving L2 learning.	Deci & Ryan (2012); Ryan & Deci (2017)	BPNS / BPNSS	Helps capture affective and social dimensions of bilingual experience.
Context of Acquisition	Environment and exposure conditions where L2 is learned (natural vs. foreign).	Krashen (1982)	Contextual classification (ESL/EFL)	Aligns with ecological and usage-based views of bilingualism.
Aptitude	Cognitive ability to learn a language efficiently (phonetic, grammatical, memory).	Carroll (1959, 1962, 1989)	MLAT	Explains variability in bilingual performance and adaptation.
Personality	Stable individual traits influencing behavior and learning.	Goldberg (1992); Jung (1971); Dörnyei (2005)	Big Five Inventory; MBTI	Adds a socio-cognitive layer to bilingualism; may explain variation in usage and control.

Note. CEFR = Common European Framework of Reference for Languages; BPNS = Basic Psychological Need Satisfaction Scale; MLAT = Modern Language Aptitude Test; MBTI = Myers-Briggs Type Indicator.

A. Individual Differences in Second Language Acquisition (SLA)

Individual differences have long been at the center of Second Language Acquisition (SLA) research, as they account for much of the variability in learners' success and learning trajectories (Dörnyei, 2005; Hummel, 2021). These variables capture both cognitive and socio-affective dimensions of language learning, providing a multidimensional view of how individuals differ in their capacity, motivation, and opportunities to acquire a second language. In this section, five of the most extensively studied individual differences in SLA—proficiency, motivation, context of acquisition, aptitude, and personality—are revisited conceptually. Each subsection outlines how these constructs have been theoretically defined and operationalized within SLA and considers their potential relevance for enhancing the description of bilingual experience. By examining these variables, this paper seeks to establish conceptual connections that may inform future empirical work in bilingualism.

(a). Proficiency

Although proficiency is a measure that has long been addressed in the field of bilingualism, here it will be examined solely from an SLA perspective. The Council of Europe (2001, 2020) published a document called "Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume" (CEFR), which is considered an international standard for describing second language ability and is used worldwide to guide the teaching of a second language (Cambridge English, 2024). It defines L2 proficiency, proposes L2 proficiency levels and descriptors, and provides key aspects for teaching and learning a second language. This widely accepted and used document is key to describing and assessing L2 proficiency anywhere in the world.

The CEFR defines proficiency as the "ability to perform communicative language activities ("can do...") while drawing upon both general and communicative language competences (linguistic, sociolinguistic, and pragmatic) and activating appropriate communicative strategies" (Cambridge English, 2024, p. 34). In this sense, the CEFR (2020) describes proficiency as a general ability, which can be developed by mastering the four skills: reading, listening, writing, and speaking. Additionally, the document classifies L2 proficiency on a six-point scale that is meant to represent progress, ranging from beginner to advanced stages. Each point on the scale represents a level, known as Common Reference Levels. These levels are A1, A2, B1, B2, C1, and C2, which can also be grouped into three broad categories: Basic user (A1 and A2), Independent user (B1 and B2), and Proficient user (C1 and C2). Having these general guidelines makes it easy for anyone involved in the field of SLA, from researchers to teachers, learners, and evaluators, to see the levels with their descriptors and make informed decisions. The standardization of the levels also allows for worldwide comparison and has helped international institutions develop standardized tests to give accurate descriptions of the proficiency level each bilingual has. Recent findings indicate that multilingual experience itself interacts with proficiency, influencing vocabulary size, lexical depth, and semantic organization across languages (Šifrar Kalan et al., 2025). These results underscore the importance of measuring proficiency as a multidimensional construct that reflects both language-specific and crosslinguistic competences.

In terms of measurement, the standardized tests that have been developed to assess L2 proficiency focus on English since it is considered today's *lingua franca*. In this context, Cambridge English, part of Cambridge University, developed Cambridge English qualifications using the CEFR as standard. This organization offers different tests that focus on specific levels. The most famous ones are the A2 Key, B1 Preliminary, B2 First, C1 Advanced, and C2 Proficiency. All these tests measure the four linguistic skills: reading, listening, speaking, and writing, plus grammar and vocabulary. The tests are subject to a constant review process and are assessed by Cambridge-certified professionals. They can be taken in certified locations around the world, making them easily accessible to anyone. Furthermore, test-takers receive a certificate with their score and the CEFR level obtained, which does not expire.

Another alternative for English learners is the TOEFL Internet-Based Test (IBT). This test also aligns with the CEFR standards and therefore assesses overall English proficiency by evaluating the four language skills, plus subskills such as grammar. Unlike the Cambridge exams, the TOEFL IBT is a single test that can measure academic English from a beginner (A1) to an advanced level (C2). As this is an internet-based test, people can take it in the comfort of their homes as long as they have an internet connection. Individuals can take the TOEFL IBT through the official webpage. Another difference from the Cambridge exams is that the certificate for this test expires after two years. This is based on the premise that languages are dynamic and change according to use.

(b). Motivation

Even though motivation is a topic that has been addressed in studies of bilingualism (mainly with heritage speakers, e.g., Jurado & García, 2018), in this paper the focus, as mentioned, will be from a Second Language Acquisition perspective. In this context, the framework that has been used to study motivation is the self-determination theory.

Self-determination theory (SDT) (Deci & Ryan, 2012; Ryan & Deci, 2017) is a psychological framework that focuses on human motivation. This theory suggests that people are inherently motivated; that is, motivation for an activity comes from within ourselves. In this sense, humans are “inherently active, intrinsically motivated, and oriented toward developing naturally through integrative processes” (Deci & Ryan, 2012, p. 417). To achieve this internal motivation, people are driven by three innate and universal psychological needs: autonomy, competence, and relatedness. Even though these three qualities are innate, they continue to develop over time and play a role in learning in general, and in language learning in particular.

Autonomy refers to the need to make one’s own decisions about what to do. In other words, it is the need to be in control of one’s own behaviors. If people develop autonomous motivation, they will not be subjected to external input and support from external influences. Autonomous motivation is self-sustaining, and therefore represents a higher quality of motivation (McEown & Oga-Baldwin, 2019). Competence refers to the need for learning and achieving mastery in the activities we perform. As a result, people may feel a sense of success because of their own abilities (Deci & Ryan, 2000). Relatedness refers to the need for a sense of belonging, to a feeling of connection to other people. This need is met if we establish meaningful and satisfying relationships with others (Furrer & Skinner, 2003; Deci & Ryan, 2012).

The core idea that SDT tries to promote is that the goal in language learning should be to develop autonomy, competence, and relatedness in the learners, so that they develop internal (or intrinsic) motivation. Teachers can “make students feel welcome and part of the language learning classroom community (relatedness), capable of successfully using the new language (competence), and that engaging with the language aligns with students’ sense of self (autonomy)” (McEown & Oga-Baldwin, 2019, p. 7). Recent studies confirm that motivational processes in multilingual learners are not static but dynamic, shaped by both intrinsic and extrinsic motives that interact across different linguistic and cultural contexts (Göksu & Louis, 2025). This evidence reinforces the need to conceptualize motivation in bilingual and SLA research as a context-dependent construct rather than a fixed trait.

Self-determination theory can be assessed through the “Basic Psychological Need Satisfaction Scales” (BPNSS). These instruments are a set of questionnaires that assess the degree to which individuals meet the three needs (autonomy, competence, and relatedness). There is a general questionnaire called “Basic Psychological Need Satisfaction Scale” (BPNS) and others specific to topics such as work and relationships. The BPNS (Deci & Ryan, 2000) has 21 items, and participants rate on a Likert scale the extent to which they agree or disagree with a given statement. This scale can be downloaded for free on the self-determination theory website.

(c). Context of Acquisition

Context of acquisition refers to the way in which bilinguals acquire their second language. In the field of Second Language Acquisition (SLA), this variable gained prominence following Stephen Krashen’s Acquisition–Learning Hypothesis (Krashen, 1982), one of several influential proposals that together shaped his theory of second language acquisition. This particular hypothesis distinguishes between two contexts of acquisition that will be discussed here.

For Krashen (1982), acquisition and learning are distinct processes. Acquisition occurs implicitly and effortlessly; it is natural and informal, similar to how children develop their first language. Learning, on the other hand, is conscious, explicit, and effortful, involving the understanding of grammar rules. The author suggested that “acquisition” is superior to “learning” for mastering a second language. He emphasized that naturalistic contexts, in which the target language functions as the community language, facilitate acquisition by providing rich and meaningful input. This view builds on Chomsky’s (1965) notion of the Language Acquisition Device (LAD), which Krashen adapted to explain adult second language acquisition. This belief also aligns with Krashen’s (1981) input hypothesis, which states that for second language learners to activate the LAD, the input must be frequent and slightly beyond their current level of competence ($i+1$). Because Krashen remains an influential figure in SLA, immersion continues to be considered—by some—the most effective method for acquiring a second language.

Although most contemporary authors do not distinguish between acquisition and learning, Krashen’s original idea created two contexts of acquisition: natural and foreign. A natural context aligns with Krashen’s concept of “acquisition” and refers to learning the L2 in a geographical location where the L2 is the dominant community language. This implies that learners should travel and immerse themselves in the target language, providing numerous natural and informal opportunities to acquire the language. A foreign context, on the other hand, is related to “learning” and typically occurs

in locations where the L1 is dominant. In this context, opportunities to use the L2 outside the classroom are scarce and limited to formal instruction. This distinction is so embedded in the field of SLA that, in the case of English learning, the acquisition process can have different labels depending on the context of acquisition. If English is acquired in a natural context, it is called English as a Second Language (ESL). However, if it is acquired in a foreign context, it is called English as a Foreign Language (EFL).

This variable has been operationalized by comparing immersion (natural context) and instruction (foreign context) and instead of applying a specific instrument to determine the context, they just consciously choose the participants based on the geographical location where they live.

(d). Language Learning Aptitude

In the field of SLA, aptitude refers to an individual's potential to learn a second language (Dörnyei, 2005; Hummel, 2016). It has been hypothesized that some learners have more language aptitude, making them better learners. Carroll (1962, 1989) is the most influential author related to this variable, and he distinguishes four constituent abilities of this individual difference: phonetic coding ability, grammatical sensitivity, rote learning ability, and inductive learning ability. Phonetic coding ability refers to the ability to identify phonemes and strings of phonemes, prosody, and intonation. It also refers to the ability to associate phonemes and graphemes. Grammatical sensitivity refers to the ability to identify the morphology and syntax in a language, such as word order, word function and form, and word meaning. Rote learning ability refers to the ability to learn associations between sounds and meanings, and between words and phrases. Inductive learning ability refers to the ability to abstract the rules and patterns of the language from new (and sometimes small) linguistic input and store them in long-term memory.

Carroll (1959, 1962, 1989) thought that these four components that constitute language aptitude could predict success in second language learning. For this reason, he created a standardized test to measure this individual difference based on these components. The test is the Modern Language Aptitude Test (MLAT). Originally, this test was designed for adults, but it is also appropriate for younger learners aged 14 and up.

The MLAT consists of five sections. The first section, Number Learning, assesses the ability to learn and recall associations between numbers and sounds in the second language. The second section, Phonetic Script, assesses the ability to associate sounds with graphemes. The third section, Spelling Clues, assesses vocabulary, more specifically, the ability to infer meaning from context. The fourth section, Words in Sentences, assesses grammatical sensitivity without using linguistic terminology. Learners must read sentences and identify the grammatical function of words. The fifth section, Paired Associates, assesses the ability to learn and recall new words in the second language.

This test takes about 60 to 90 minutes to complete, and each section has a specific time limit of 10-15 minutes. Currently, this test is owned by the Language Learning and Testing Foundation, a non-profit entity based in the United States founded by Dr. Charles Stansfield. This test is only available for purchase by government agencies, missionary groups, and licensed clinical psychologists.

(e). Personality

Personality has been defined as "all the attributes that characterize a unique individual" (Dörnyei, 2005, p. 11). These attributes tend to follow a pattern; that is, personality is stable and does not change according to the situation in which a person engages. Unlike temperament or mood, personality shows consistency in the way a person behaves, regardless of the situation they are in. Several attempts have been made to classify different personality types. The most famous ones are the 'Big Five' model (Goldberg, 1992; McCrae & Costa, 2003) and Jung's (1971) theory of personality types, which gave rise to the Myers-Briggs Type Indicator (MBTI) (Briggs & Myers, 1976).

The Big Five Model, unlike the MBTI, does not have a single author because it was developed over several decades with contributions from different researchers (e.g., Eysenck et al., 1985; Goldberg, 1992; McCrae & Costa, 2003).

The Big Five Model identifies five broad dimensions of personality that are also called OCEAN because of its acronym. The first one is "Openness to experience" (O), which refers to the tendency to be imaginative, curious, flexible, creative, original, and open-minded. The second dimension is "Conscientiousness" (C), which refers to the tendency to be systematic, meticulous, efficient, organized, responsible, and hard-working. The third dimension is "Extraversion-Introversion" (E), which refers to the tendency to be either extroverted (sociable, active, assertive, talkative, outgoing) or introverted (passive, quiet, reserved, withdrawn). The fourth dimension is "Agreeableness" (A), which refers to the tendency to be compassionate, good-natured, cooperative, friendly, kind, trusting, and generous. The fifth and last dimension is "Neuroticism-Emotional Stability" (N), which refers to the tendency to experience anxiety, depression, moodiness, and emotional instability.

Jung's (1971) theory of personality types, which gave rise to the Myers-Briggs Type Indicator (MBTI) (Myers & Briggs, 1976), consists of dichotomous bipolar types: extraversion (E)-introversion (I), sensing (S)-intuiting (N), and thinking (T)-feeling (F). Briggs and Myers (1976) added a fourth dichotomy: judging (J)-perceiving (P). The first dimension refers to how individuals prefer to focus their attention and spend their energy. While extroverts are active, sociable, talkative, and energetic, introverts are reserved and withdrawn. The second dimension refers to how individuals prefer to collect information. While sensors prefer to focus on tangible objects, intuitives prefer to look at patterns. The third dimension refers to how individuals prefer to make decisions. While thinkers make decisions based on logic, feelers prefer to prioritize personal values. The last dimension refers to how individuals prefer to interact with the external world.

While judgers prefer structure and organization, perceivers are more flexible and spontaneous.

The MBTI is an instrument designed to categorize personality types, where each individual's personality is formed by a combination of four letters, one from each of the dichotomous options shown above. For example, someone with an ESTJ personality is more inclined to being extroverted, sensing, thinking, and judging.

In the field of SLA, the most researched personality aspect has been the extraversion–introversion dimension (Kim, 2022; Zafar & Meenakshi, 2012), present in both models (The Big Five and the MBTI). It has been hypothesized that extroverts would be better language learners than introverts because they are outgoing and energetic, are willing to make mistakes, and engage in conversations inside and outside of the classroom. However, the results are still mixed. Even though the advantage of extroverts has been documented in speaking skills, they have not been consistent (see Octaviani et al., 2023, for a literature review). Some studies have shown that introverts may be better at classroom learning (Zafar et al., 2017), especially because introverts benefit more from receptive skills (Araghi et al., 2011). Other studies, however, did not find significant differences in overall proficiency between the two personality types (Alagić, 2022; Chen & Tsai, 2015).

III. DISCUSSION

A. *Transdisciplinary Integration Between Bilingualism and SLA*

The present paper sought to establish a dialogue between the fields of Bilingualism and Second Language Acquisition (SLA). While these two areas have historically evolved along parallel paths, they share a fundamental object of study: the individual who learns, uses, and adapts through more than one language. Bilingualism has focused on understanding how experience with multiple languages shapes cognition and neural organization, whereas SLA has primarily examined the mechanisms that underlie language learning and the variables that explain interindividual variation. What emerges from bringing them together is not merely an interdisciplinary exchange, but a truly transdisciplinary synthesis—one that aims to build a common conceptual ground from which both fields can rethink their assumptions.

In this sense, bilingualism can benefit from SLA's long tradition of theorizing individual differences and developing reliable instruments to measure them, while SLA can, in turn, benefit from bilingualism's ecological and dynamic view of language use. This reciprocal integration allows for a richer and more flexible understanding of the bilingual individual—not as a static category defined by linguistic boundaries, but as a learner embedded in continuous adaptation. This perspective aligns with recent syntheses highlighting emerging directions in bilingualism research, including social interactional dynamics, neural plasticity, and the lifelong adaptability of bilingual cognition (MacWhinney & Gao, 2025). Such developments further emphasize the need for integrative frameworks that connect cognitive, experiential, and social dimensions of bilingual experience.

B. *Epistemological Shifts: From Learners to Learning Individuals*

One of the most significant outcomes of this transdisciplinary dialogue lies in the epistemological shift it implies. Traditionally, SLA has conceptualized its subjects as language learners—individuals engaged in acquiring an additional language, often in formal or semi-formal contexts. Bilingualism, by contrast, has tended to view its participants as language users—individuals who already navigate multiple linguistic systems. By connecting these two perspectives, a new conceptual category emerges: the learning individual, whose language experience reflects a continuum of acquisition, use, and adaptation.

From this viewpoint, bilingualism is not a state, but a process—a dynamic expression of learning in context. Similarly, SLA is not confined to early stages of learning, but extends into the adaptive, integrative, and identity-related dimensions that characterize bilingual life. This reconceptualization challenges the dichotomy between learning and using a language, suggesting instead that both belong to the same spectrum of human learning and cognition.

This shift also contributes to redefining the epistemological boundaries of both fields. It positions language learning and use within a broader understanding of learning as adaptation, where cognitive, affective, and contextual factors constantly interact. By acknowledging that bilinguals are not simply speakers of two languages, but learners for life, this approach situates bilingualism within the wider framework of human development and plasticity.

C. *Methodological Convergence*

Beyond the theoretical implications, this synthesis has methodological consequences for how researchers describe and measure bilingual experience. SLA has long relied on theoretically grounded constructs—such as proficiency, motivation, aptitude, and personality—to explain variation in language learning outcomes. The systematic assessment of these variables, through validated instruments like the CEFR, BPNS, MLAT, or personality inventories, offers a structured way to capture dimensions of individual difference that have often been overlooked in bilingualism research. In line with this, recent theoretical work has called for reconceptualizing individual differences as dynamic, contextually embedded, and mutually interacting systems, encouraging methodological innovation and more holistic approaches to data integration (Yang & Gao, 2025). Such perspectives reinforce the idea that future research on bilingualism should adopt designs capable of capturing the evolving and interconnected nature of learner variability.

At the same time, bilingualism offers SLA valuable insights into how learning unfolds in naturalistic, real-world contexts. Its emphasis on ecological validity, contextual diversity, and lived experience reminds us that learning does not

occur in isolation, but within specific social, cultural, and affective environments. Bringing these methodological traditions together opens the possibility for a more comprehensive framework—one that could be called an Ecological Individual Differences Framework for Bilingualism. Such an approach would integrate cognitive measures (like aptitude or proficiency) with contextual and motivational data, capturing the full complexity of bilingual experience.

This kind of methodological cross-fertilization would also strengthen both disciplines' capacity for replication and synthesis. Bilingualism studies could gain precision and reliability, while SLA research could expand its ecological reach. In doing so, both fields would converge toward a shared goal: developing integrative, multidimensional profiles of language learners and users.

D. Alignment With the Science of Learning

This transdisciplinary integration aligns naturally with the broader principles of the Science of Learning, an emerging field that seeks to understand learning as a complex, adaptive, and context-dependent process. From this perspective, language learning—and bilingualism as its natural extension—are not isolated phenomena, but manifestations of general learning mechanisms shaped by cognition, emotion, and environment. The Science of Learning encourages precisely the kind of dialogue proposed here: one that transcends disciplinary boundaries to build models that reflect the richness of human experience.

Within this framework, both SLA and bilingualism can be understood as complementary lenses for studying how humans learn, adapt, and create meaning through language. SLA contributes insights into individual variability and the mechanisms of acquisition, while bilingualism contributes an understanding of how experience reshapes those mechanisms over time. Together, they advance a vision of language learning that is dynamic, developmental, and deeply human.

This view also resonates with the efforts of research initiatives such as Chile's Millenium Nucleus for the Science of Learning (MiNSoL) (MiNSoL, 2025), which promote collaboration between cognitive, linguistic, and educational sciences to generate integrative accounts of learning. By situating bilingualism and SLA within this transdisciplinary landscape, we can move beyond describing language as a system and toward understanding it as a vehicle of adaptation, identity, and continuous learning. Ultimately, such an approach does not merely connect two fields—it expands the very meaning of what it means to learn.

In summary, the integration of SLA and bilingualism represents more than a conceptual bridge; it is a reimagining of language learning itself. Through transdisciplinary collaboration, epistemological rethinking, and methodological renewal—anchored in the Science of Learning—both fields can contribute to a shared understanding of the bilingual individual as an active, adaptive learner across the lifespan.

IV. CONCLUSION

The fields of Bilingualism and Second Language Acquisition (SLA) share a common object of study—the individual who learns and uses more than one language—but they have traditionally approached it from distinct disciplinary perspectives. While bilingualism has drawn primarily on cognitive science and neuroscience to explain how experience shapes the mind and brain, SLA has evolved within psychology, linguistics, and education to explain the mechanisms and variability underlying language learning. These different trajectories have produced rich, yet parallel, bodies of knowledge that have only recently begun to converge.

This paper proposed a conceptual bridge between the two fields, illustrating how SLA-derived constructs—such as proficiency, motivation, context of acquisition, aptitude, and personality—can contribute to a more systematic and multidimensional understanding of bilingual experience. By revisiting these constructs through an SLA lens, the paper highlights new ways of operationalizing individual differences in bilingualism research and invites bilingualism scholars to consider well-established tools that capture cognitive and affective variability.

Beyond this disciplinary connection, the paper aligns with the broader principles of the *Science of Learning*, which calls for integrative, context-sensitive approaches to understanding how humans learn and adapt. From this perspective, uniting bilingualism and SLA represents more than theoretical alignment; it is a transdisciplinary effort to understand language learning as a lifelong, adaptive process that reflects the dynamic interplay between cognition, experience, and environment.

Ultimately, bridging these fields opens the door to more precise, ecologically valid, and human-centered research. It encourages us, as researchers, to move beyond categorical labels and instead describe bilinguals as learning individuals whose linguistic and cognitive trajectories are shaped by diverse experiences. In doing so, we not only refine how we study bilingualism but also contribute to a larger conversation about how learning—through language—defines the human experience.

FUNDING

The author acknowledges the support of the fellowship granted by the Chilean National Agency (ANID), Beca Doctorado Nacional – 21201324, to Mackarena Kartsevski and the support of the Millenium Science Initiative through grant ANID – MILENIO - NCS2022_026.

ACKNOWLEDGMENTS

The author would like to express her sincere gratitude to Dr. Roberto Ferreira, Director of the Millennium Nucleus for the Science of Learning (MiNSoL), for his insightful suggestions on the methodological organization and overall structure of this paper.

REFERENCES

- [1] Alagić, M. (2022). Extroversion and introversion in second language acquisition. *MAP Education and Humanities*, 22-29. <https://doi.org/10.53880/2744-2373.2022.2.2.22>
- [2] Araghi, M., Fam, A., & Ziaei, E. (2011). The effect of personality on the English as foreign language learners' performance on listening comprehension: extroverts versus introverts. *International Journal of Applied Linguistic Studies*, 1(1), 11–16.
- [3] Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(6), 1022–1047. <https://doi.org/10.1037/xlm0000770>
- [4] Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, 143(3), 233–262. <https://doi.org/10.1037/bul0000099>
- [5] Bialystok, E., & Craik, F. I. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic bulletin & review*, 29(4), 1246–1269.
- [6] Bowen, A. E. J., Ferreira, R. A., Tolmie, A., Thomas, M. S. C., Borst, G., & Van Herwegen, J. (2025). International perspectives on gaps and solutions for integrating research evidence into classroom practices. *NPJ Science of Learning*, 10(1), 79.
- [7] Bratlie, S. S., Grøver, V., Lekhal, R., Chen, S., & Rydland, V. (2025). Home literacy environment, language use, and proficiency: bilingual profiles in young learners. *Journal of Applied Developmental Psychology*, 96, 101728.
- [8] Briggs, K. C., & Myers, I. B. (1976). *Myers-Briggs Type Indicator*. Consulting Psychologists Press.
- [9] Cambridge English. (2024). *International language standards*. Retrieved January 10, 2025, from <https://www.cambridgeenglish.org/exams-and-tests/cefr/>
- [10] Carroll, J. B. (1959). Use of the Modern Language Aptitude Test in secondary schools. *Yearbook of the National Council on Measurements Used in Education*, 16, 155–159.
- [11] Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training, research and education* (pp. 87–136). University of Pittsburgh Press.
- [12] Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26–31.
- [13] Chen, W., & Tsai, M. (2015). A preliminary examination of the relationships among extroversion-introversion personality, learning strategies and English proficiency. *Journal of Chaoyang University*, 13(1), 33-56.
- [14] Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- [15] Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.
- [16] Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. <https://www.coe.int/lang-cefr>
- [17] De Bruin, A. (2019). Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences. *Behavioral Sciences*, 9(3), 33. <https://doi.org/10.3390/bs9030033>
- [18] Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- [19] Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange et al. (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 416–436). Sage. <https://doi.org/10.4135/9781446249215.n21>
- [20] Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.
- [21] Dumont, H., Istance, D., & Benavides, F. (Eds.). (2010). *The Nature of Learning: Using Research to Inspire Practice, Educational Research and Innovation*. OECD Publishing, Paris, <https://doi.org/10.1787/9789264086487-en>
- [22] Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, 6(1), 21–29. [https://doi.org/10.1016/0191-8869\(85\)90026-1](https://doi.org/10.1016/0191-8869(85)90026-1)
- [23] Ferreira, R. A., & Rodríguez, C. (2022). Effect of a science of learning course on beliefs in neuromyths and neuroscience literacy. *Brain Sciences*, 12(7), Article 811.
- [24] Fischer, K. W., Goswami, U., & Geake, J. (2010). The future of educational neuroscience. *Mind, Brain, and Education*, 4(2), 68–80. <https://doi.org/10.1111/j.1751-228X.2010.01086.x>
- [25] Furrer, C., & Skinner, E. A. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95(1), 148–162. <https://doi.org/10.1037/0022-0663.95.1.148>
- [26] Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26–42.
- [27] Göksu, A., & Louis, V. (2025). Motivational dynamics in a multilingual context: University students' perspectives on LOTE learning. *Behavioral Sciences*, 15(7), 931. <https://doi.org/10.3390/bs15070931>
- [28] Hummel, K. M. (2021). *Introducing second language acquisition: Perspectives and practices* (2nd ed.). Routledge.
- [29] Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*, 1(1), 3–10. <https://doi.org/10.1111/j.1751-228X.2007.00004.x>
- [30] Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS review*, 10(1), 18-26.
- [31] Jung, C. G. (1971). *Psychological types*. Princeton University Press. (Original work published 1923)
- [32] Jurado, B. C., & García, C. M. (2018). Students' attitude and motivation in bilingual education. *International Journal of Educational Psychology*, 7(3), 317–342. <https://doi.org/10.17583/ijep.2018.3488>

- [33] Kim, Y. R. (2022). Extraversion vs. introversion: Comparative and contrastive analysis of empirical studies in second language learning and teaching. *English Teaching*, 17(2), 213–230.
- [34] Krashen, S. (1981). Second language acquisition. *Second Language Learning*, 3(7), 19–39.
- [35] Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- [36] Kroll, J. F., & Rossi, E. (2023). Models and metaphors. In G. Luk, J. A. E. Anderson, & J. G. Grundy (Eds.), *Understanding language and cognition through bilingualism* (pp. 210–234). John Benjamins.
- [37] Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.
- [38] López, B. G., Luque, A., & Piña-Watson, B. (2021). Context, intersectionality, and resilience: Moving toward a more holistic study of bilingualism in cognitive science. *Cultural Diversity & Ethnic Minority Psychology*, 29(1), 24–35. <https://doi.org/10.1037/cdp0000440>
- [39] Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605–621. <https://doi.org/10.1080/20445911.2013.795574>
- [40] MacWhinney, B., & Gao, Z. (2025). Six advances in research on bilingualism. *Brain and Language*, 268, 105614.
- [41] Meir, N. (2023). Individual differences in bilingual child language acquisition: a plunge into a complex and dynamic network. *Journal of Child Language*, 50(4), 827–831. <https://doi.org/10.1017/S0305000923000016>
- [42] McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). Guilford Press.
- [43] McEown, M. S., & Oga-Baldwin, W. Q. (2019). Self-determination for all language learners: New applications for formal language education. *System*, 86, 102124. <https://doi.org/10.1016/j.system.2019.102124>
- [44] Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science*, 325(5938), 284–288.
- [45] Millennium Nucleus for the Science of Learning (MiNSoL). (2025). *About MiNSoL*. Retrieved November 25, 2025, from <https://minsol.cl>
- [46] Octaviani, S. N., Jumariati, J., Elyani, E. P., & Nasrullah, N. (2023). Learner's personalities and language learning: A literature review. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 11(1), 218–231. <https://doi.org/10.24256/ideas.v11i1.3330>
- [47] OECD. (2018). *The Future of Education and Skills: Education 2030 – The OECD Learning Framework 2030*. OECD Publishing.
- [48] Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, 18(1), 1–25. <https://doi.org/10.1017/S0272263100014662>
- [49] Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, 50(4), 793–817.
- [50] Peal, E., & Lambert, W. (1962). The relation of bilingualism to intelligence. *Psychological Monographs*, 76(27), 1–23. <https://doi.org/10.1037/h0093840>
- [51] Pliatsikas, C. (2023). Bilingualism and brain structure. In G. Luk, J. A. E. Anderson, & J. G. Grundy (Eds.), *Understanding language and cognition through bilingualism* (pp. 301–324). John Benjamins.
- [52] Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.
- [53] Šifrar Kalan, M., Muñoz-Basols, J., Robles-García, P., Strawbridge, T., & Sánchez-Gutiérrez, C. (2025). The impact of multilingualism and proficiency on L2 vocabulary knowledge: contrasting high and low multilinguals. *International Journal of Multilingualism*, 22(2), 433–456.
- [54] Swain, M., & Cummins, J. (1979). Bilingualism, cognitive functioning and education. *Language Teaching*, 12(1), 4–18. <https://doi.org/10.1017/S026144480000415X>
- [55] Thomas, M. S. C., & Ansari, D. (2020). Educational neuroscience: Why is neuroscience relevant to education? In M. S. C. Thomas, D. Mareschal, & I. Dumontheil (Eds.), *Educational neuroscience: Development across the life span* (pp. 3–22). Routledge/Taylor & Francis Group. <https://doi.org/10.4324/9781003016830-2>
- [56] Tokuhama-Espinosa, T. (2011). *The new science of teaching and learning: using the best of mind, brain, and education science in the classroom*. Teachers College Press.
- [57] van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, 9(3), 27. <https://doi.org/10.3390/bs9030027>
- [58] Yang, W., & Gao, X. (2025). Envisioning New Directions in Research on Individual Differences. *Studies in Second Language Learning and Teaching*, 15(2), 427–443.
- [59] Yow, W. Q., & Li, X. (2024). Role of bilingual experience in children's context-sensitive selective trust strategies. *Bilingualism: Language and cognition*, 27(1), 95–106.
- [60] Zafar, S., & Meenakshi, K. (2012). A study on the relationship between extroversion–introversion and risk-taking in the context of second language acquisition. *International Journal of Research Studies in Language Learning*, 1(1), 33–40. <https://doi.org/10.5861/ijrsl.2012.135>
- [61] Zafar, S., Khan, Z., & Meenakshi, K. (2017). Extraversion–introversion tendencies and their relationship with ESL proficiency: A study of Chinese students in Vellore, India. *Pertanika Journal of Social Sciences & Humanities*, 25(2), 687–703.

Mackarena Kartsevski received the B.A. degree in English Language and Literature and the B.A. degree in Education, both from the Pontificia Universidad Católica de Valparaíso, Chile, with *Cum Laude* distinction. She obtained the M.A. degree in Applied Linguistics from the Universidad de Concepción, Chile, in 2019, and the Ph.D. degree in Linguistics from the Pontificia Universidad Católica de Valparaíso, Chile, in 2025, *Summa Cum Laude*. As part of her doctoral training, she completed a research internship at the Bilingualism, Mind & Brain Lab, directed by Dr. Judith Kroll, at the University of California, Irvine.

She is currently an Academic at the Universidad de Talca, Chile, and a Researcher at the Millennium Nucleus for the Science of Learning (MiNSoL). Her previous academic experience includes teaching and research positions at the Pontificia Universidad Católica de Valparaíso. She has collaborated with the California branch of Bilingualism Matters and presented her work at international conferences such as the Psychonomic Society and the University of Iowa ROLE Symposium. Her publications appear in journals such as *Literatura y Lingüística* and *Revista de Estudios y Experiencias en Educación*.

Dr. Kartsevski's research explores bilingualism, second language acquisition, and how learning and development unfold across the lifespan. She is a member of the Psychonomic Society, the Chilean Linguistics Society (SOCHIL), and the National Society of Language Teachers in Higher Education (SONAPLES). Email: mackarena.kartsevski@utalca.cl; ORCID iD: 0000-0003-0228-3836.

Why Are Students Failing Listening? Investigating Reasons for the Low Performance of Omani General Foundation Program Students in English Listening Exams

Fakhriya H. Al-Rashdi

Preparatory Studies Centre, University of Technology and Applied Sciences, Muscat, Sultanate of Oman

Abstract—This research paper investigates the reasons behind the low performance of students in the listening component of English examinations within the General Foundation Program (GFP) at an Omani public university. Recognizing a consistent trend of weak listening outcomes across levels and branches, the study aims to identify the exam sections and question types that pose the greatest challenges to learners. Using data from the university's largest branch, the study employs Microsoft Office Excel for performance analysis and item analysis to determine the most problematic questions. The listening examinations, which are standardized across all branches, are designed to align with students' proficiency levels and include multiple-choice, true or false, short-answer, and note-completion tasks. The findings indicate that students experience particular difficulty with cloze and note-completion tasks due to the demands of precise listening, linguistic decoding, and accurate spelling or transcription of words and numbers. The study also highlights inherent challenges of listening comprehension, including real-time processing, exposure to varied accents, differences in speech rate, and the difficulty of converting auditory input into written responses. These findings are relevant not only to the context examined in this study but also to other educational settings where listening skills are taught and assessed. The study underscores the need for targeted pedagogical interventions to improve listening proficiency and assessment performance, thereby supporting students' overall English competence and academic success in programs where English is the medium of instruction.

Index Terms—listening comprehension, EFL learners, language assessment, foundation programs, English listening

I. INTRODUCTION

In the Sultanate of Oman, the General Foundation Program (GFP) serves as a preparatory or bridging program for students before they enter a degree program. In other words, the GFP functions as a bridge between secondary and higher education, and only upon successful completion may students progress directly into their chosen specializations. The GFP is required for students who have not yet achieved the necessary level of proficiency in the four core areas—English, Mathematics, Computing, and General Study Skills—to pursue their studies in higher education. According to the Oman Authority for Academic Accreditation and Quality Assurance of Education (OAAAQA), the entity responsible for quality assurance and accreditation in the Sultanate, ‘the GFP is a compulsory entrance qualification for Omani degree programs’ and ‘is designed to prepare students for their post-secondary and higher education studies’ (OAS, p. 6). The OAS further states that the GFP aims to help students ‘attain the prescribed student learning outcomes in at least four areas: English, Mathematics, Computing and General Study Skills’ (OAS, p. 4). As far as the English component is concerned, the aim of the program is to enhance students’ academic English proficiency, particularly in programs where English is the medium of instruction. Students are expected to achieve a level equivalent to IELTS Band 5 or higher upon completion of the English program to qualify for entry into their degree programs.

This Omani Public University (OPU) is a branch-based Higher Education Institution with 11 branches strategically dispersed throughout the Sultanate of Oman. It is the largest university in the country. The GFP is managed by the Preparatory Studies Centre (PSC), which oversees development, delivery, and review of the program across all branches. Each branch has its own PSC, which is responsible for implementing, monitoring, and reviewing the GFP locally, while all 11 centers report directly to the main PSC at the university headquarters.

All Preparatory Studies Centres (PSCs) across the branches follow a common GFP structure. Teachers in English, Mathematics, and Computing implement the same syllabi, course outlines, assessment system, and teaching materials. In the GFP English Program, which is the focus of this study, there were four levels of study at the time of data collection. Upon taking a placement test, students were placed into one of four levels: Level 1 (CEFR A1), Level 2 (CEFR A2), Level 3 (CEFR B1), or Level 4 (CEFR B2, lower range). Students who achieved an overall IELTS score of 5.0 through an official certificate were exempt from the English program. Across all levels, the four language skills—Reading, Writing, Listening, and Speaking—were given equal importance in terms of teaching hours, materials, and assessment. Although

the program has since been restructured into three General English courses, namely General English 1 (CEFR A2), General English 2 (CEFR B1), and General English 3 (CEFR B2, lower range), the Listening examinations and their assessment procedures have remained unchanged.

As part of the OPU GFP English assessment system, students across all branches sit the same Midterm Examination (MTE) and Final Examination (FE). The MTE, which constitutes 30% of the continuous assessment, evaluates three skills—Reading, Writing, and Listening—with 10% allocated to each skill. The FE, which accounts for 50% of the overall course grade, assesses all four skills equally (12.5% each). While the MTE mirrors the exam style of the FE, it assesses only the learning outcomes taught up to the point of the midterm, whereas the FE is comprehensive and assesses any learning outcome covered during the semester.

Following the implementation of the common GFP English program, it has been consistently observed that students across all branches perform less well on Listening exams across all levels. This study was therefore undertaken to investigate the possible reasons behind this trend. Specifically, the study examined the three sections of the Listening exams and identified which section posed the greatest difficulty for students. It also explored the most challenging question types.

II. LITERATURE REVIEW

English is recognized as a global lingua franca and is spoken by millions of people worldwide. According to Statista, based on data reported for 2022, approximately 200 million people speak English as a native language, with over one billion using it as a second or foreign language (Statista, 2023). People learn English for many reasons, including its role in facilitating communication, supporting education and career advancement, and providing access to information and cultural resources.

In Oman, English is the only official foreign language and is often considered a lingua franca as well (Al-Issa, 2014). It is taught as a mandatory subject throughout the formal education system, beginning in Grade 1 with basic skills such as vocabulary, simple sentence structures, and basic communication. English remains compulsory through intermediate and secondary education, following a structured curriculum set by the Ministry of Education. This curriculum emphasizes the development of reading, writing, listening, and speaking skills, as well as grammar and vocabulary. Public schools use the same prescribed English language textbooks aligned with the Ministry's curriculum, and instruction commonly incorporates a mix of traditional teaching methods, interactive activities, group work, and language games.

Language acquisition refers to the process by which individuals acquire the ability to understand and use language. Children naturally acquire their mother tongue through exposure and interaction, and researchers have attributed this process to a combination of innate biological factors, environmental exposure, social interaction, and cognitive development (Herschensohn & Young-Scholten, 2013). In contrast, second language (L2) acquisition does not typically occur in the same way and often requires more time. A variety of factors influence the language acquisition journey of ESL learners, including age, previous language experience, exposure to the target language, and the learning environment (Gass & Schachter, 1989). Moreover, L2 learners do not necessarily master the four language skills—listening, speaking, reading, and writing—uniformly. Differences in learning styles, preferences, and strengths contribute to varied levels of difficulty across skills. The learning environment and instructional methods also play significant roles in shaping learners' experiences (Herschensohn & Young-Scholten, 2013).

An important question arises: Is there a relationship between language acquisition and structured learning? Research suggests that structured learning can greatly benefit foreign language acquisition. A well-designed curriculum offers learners a systematic framework to develop their skills (Ibbotson, 2022; Krashen, 1982). However, the effectiveness of a common English curriculum depends on several factors, including curriculum design, instructional quality, learner characteristics, and the broader learning environment. Although standardized curricula provide a consistent structure, individual differences and contextual factors inevitably influence outcomes. Learners perceive the difficulty of English language skills differently; what one learner finds challenging may not be difficult for another. Nevertheless, a substantial body of research highlights listening as one of the most challenging skills for L2 learners (Buck, 2001; Vandergrift, 2007; Graham, 2006).

Listening poses significant difficulties due to intrinsic features of spoken language. Unlike reading, listening requires real-time processing, leaving no opportunity to review or reflect (Buck, 2001). Variability in accents, pronunciation, and speaking speeds adds to the complexity, as learners must adjust to linguistic nuances and dialectal differences (Vandergrift, 2007). Recent theoretical work further emphasizes the cognitive and real-time processing demands that distinguish listening from other language skills (Field, 2025). Spoken English also contains colloquial expressions, idioms, and reduced forms that are not commonly found in written language, further complicating comprehension (Graham, 2006). Background noise and overlapping speech in natural settings can hinder understanding. Collectively, these factors make listening one of the most demanding skills for L2 learners, requiring strong linguistic knowledge, decoding ability, and inferencing skills.

Since this study focuses on the listening performance of Omani students in OPU GFP Listening examinations, it is important to explore challenges commonly faced by Arabic learners of English. Hasan (2000) identifies pronunciation as a major difficulty due to substantial differences between English and Arabic phonological systems. English consonant clusters and vowel contrasts pose notable challenges. Grammar also presents difficulties, given the complexity and

irregularity of English structures compared with the more systematic nature of Arabic grammar. Vocabulary acquisition, too, is demanding due to the substantial lexical differences between the two languages. Hasan's research highlights pronunciation, grammar, and vocabulary as interconnected challenges that collectively shape the Arabic learner's linguistic experience.

Arabic L2 learners also encounter difficulties in the IELTS Listening Module. Phonological disparities between English and Arabic create challenges in distinguishing and reproducing English sounds (Al-Shehri, 2017). Grammatical differences between the two languages lead to difficulties interpreting complex English sentences (Elsheikh, 2002; Zughoul, 2002). Limited vocabulary further hinders comprehension (Al-Obaidi, 2019). These combined factors contribute to the challenges Arabic speakers face in achieving proficiency in IELTS Listening.

Students also struggle with specific question types in listening assessments. Research consistently identifies cloze exercises as the most challenging. These tasks require test takers to fill in blanks in a paraphrased passage based on an audio input. The difficulty stems from the need for precise listening, accurate spelling, and the ability to recognize paraphrased content. Studies by Buck (2001) and Goh (2000) highlight the cognitive load associated with these tasks, which require simultaneous listening, processing, and writing. Field (2008) further notes that cloze tasks demand understanding of context and prediction of linguistic structures, increasing the cognitive burden on learners. Recent listening assessment and SLA-oriented research confirms that gap-fill and cloze-type tasks remain among the most cognitively demanding listening formats for EFL learners due to their transcription, spelling, and processing demands (Aryadoust et al., 2021; Aryadoust et al., 2025).

Other question types also pose challenges. Multiple-choice questions (MCQs), which often test details such as dates, times, and names, can overload learners with dense information. Roediger and Marsh (2005) argue that MCQs increase cognitive load by requiring evaluation of several options simultaneously, while Bridgeman and Rock (1993) note that they may emphasize recognition over comprehension. Recent research further demonstrates that multiple-choice formats in listening assessments can influence learner performance depending on question design and distractor characteristics, particularly when lexical or numerical overlap increases cognitive demands (O'Grady, 2021; Ockey, 2025).

Sentence completion tasks require vocabulary knowledge, grammatical awareness, and inferencing skills (Frase, 1968; Brown & Hudson, 1998). Their open-ended nature can induce anxiety. Short-answer questions demand accurate spelling, specific information recall, and effective note-taking abilities, which are difficult for many L2 learners (Kiewra, 1989). Contemporary research on academic listening continues to show that note-taking ability plays a critical role in listening comprehension and assessment performance (Siegel & Wang, 2025).

Effective listening requires a wide range of skills and strategies. Learners must engage actively, employ bottom-up and top-down processing, use predictive strategies, apply varied listening techniques, manage vocabulary and grammar demands, develop note-taking skills, interpret contextual clues, understand different accents, and reflect regularly on their performance (Field, 2008; Vandergrift, 2007; Graham, 2006). These combined demands illustrate why listening is one of the most complex skills to acquire in a second language.

III. METHODOLOGY

A. Research Questions

This study addressed the following research questions:

1. What are the possible reasons for the low achievement of OPU GFP students in their listening exams?
2. Which exam section is the most problematic to GFP students?
3. What type of questions do students struggle with the most and why?

B. Participants

The study examined the performance of OPU General Foundation Program (GFP) students across the four proficiency levels offered at the time of data collection. Listening examination results from four semesters, Fall and Spring of AY 2021–2022 and Fall and Spring of AY 2022–2023, were analyzed to identify patterns and trends in students' performance over time. Although the GFP English program has since been restructured into three General English courses (GE1, GE2, and GE3), the Listening examinations and assessment procedures remained unchanged. Data were obtained from the largest branch of OPU, which enrolls the highest number of GFP students and therefore provides the most comprehensive and analyzable dataset.

C. Instruments

The main instrument used in this study was the standardized Final Examination (FE) Listening paper, administered at four GFP levels. Each Listening exam consisted of 25 questions divided into three sections, which varied in the number of speakers, the length of the listening scripts, the task types, and the cognitive demands placed on test takers. The examinations were designed according to level-specific testing specifications that defined:

- the learning outcomes assessed in each section
- the number of speakers involved
- the length of the listening texts
- the topics / themes covered

- whether the text was a monologue or a conversation
- the test duration, including pause time
- the task types included (e.g., multiple choice, true/false, note completion, short answer questions)
- the number of times the recordings were played
- the sources of the scripts
- the text difficulty aligned with CEFR levels (A1, A2, B1, B2)

All Listening examinations were administered uniformly across OPU branches and were centrally marked using an answer key and a double-marking process to ensure consistency and fairness. Table 1 presents a comparison of the Listening examinations across Levels 1–4.

TABLE 1
A COMPARISON BETWEEN LEVELS 1-4 LISTENING EXAMINATIONS

Level	Section 1			Section 2		Section 3	
	Dialogue between 2 people			Conversation between 3 people		Monologue	
	Total Questions: L1–L2 = 10 L3–L4 = 9			Total Questions: L1–L2 = 5 L3–L4 = 6		Total Questions: All levels = 10	
	Task types			Task types		Task types	
One	True / False 3	MCQ* 4	SAQ** 3	Table / Form completion 5	True / False 4	MCQ 4	Sentence Completion 2
Two	True / False 3	MCQ 4	SAQ 3	Table / Form completion 5	MCQ 5		Sentence Completion/ Note Completion 5
Three	True / False 3	MCQ 3	SAQ 3	Table completion 6	True / False 3	SAQ 2	Note completion 5
Four	MCQ 3	SAQ 3	True / False 3	Table completion 6	True / False 3	SAQ 2	Note completion 5

Note. *MCQ: Multiple Choice Question; **SAQ: Short-answer Question

In all levels, the Listening exam consisted of a question paper and an answer sheet. The recordings were played twice. While listening, students wrote their answers on the question paper and were then given six minutes at the end of the exam to transfer their responses to the answer sheet. For items requiring written responses such as names, nationalities, words, prices, and numbers, students were required to provide grammatically correct and correctly spelled answers. Less familiar proper names, as well as some country and city names, were spelled out in the recordings.

Table 2 below shows the number and proportion of questions requiring written responses at each level. The proportion of written-response items increases across the levels, with 10 questions (40%) in Level 1, 13 questions (52%) in Level 2, and 16 questions (64%) in Levels 3 and 4.

TABLE 2
QUESTIONS REQUIRING WRITTEN RESPONSES

Level	Section 1	Section 2	Section 3		Total
One	SAQ 3	Table / Form completion 5	Sentence Completion 2		10
Two	SAQ 3	Table / Form completion 5	Sentence/ Note Completion 5		13
Three	SAQ 3	Table completion 6	SAQ 2	Note completion 5	16
Four	SAQ 3	Table completion 6	SAQ 2	Note completion 5	16

D. Data Collection Procedures

Data collection was carried out in two phases, each using a specific analytical approach to examine students' performance in the Listening examinations.

Phase One: Section-Level Performance Analysis

Final Examination (FE) Listening results from four semesters, Fall and Spring of AY 2021–2022 and Fall and Spring of AY 2022–2023, were collected and entered into an Excel-based analytical tool. For each level, the tool generated:

- the total number of correct answers per section
- patterns of section-level difficulty
- comparisons across semesters and academic years

This phase aimed to identify which section of the Listening exam consistently produced the lowest performance across Levels 1 to 4.

Phase Two: Item-Level Analysis

In the second phase, item-by-item analysis was carried out using sampled Listening Exam papers from Spring AY 2022–2023. A random sampling procedure was used in which half of the groups at each level were selected for analysis.

Two structured forms were used during this process:

- Teacher Form, used for recording every correct answer for each student response
- Supervisor Form, used to verify the accuracy and completeness of the teachers' entries

The analysis was conducted by four teams, with each team assigned to one level of the Listening exam. Each team worked under the guidance of a supervisor. The analysis took place in a designated computer lab where teachers worked in pairs; one teacher checked each question paper and the other entered the corresponding correct responses into the form. The form automatically calculated the total number of correct answers for every item.

At the end of each session, teachers submitted their completed forms to the supervisor, who checked the entries using the Supervisor Form and then compiled and forwarded the verified datasets to the researcher for further analysis. This phase aimed to identify the most difficult questions and the specific sources of difficulty, including challenges related to spelling, recognition of paraphrased language, numerical transcription, inference, and idiomatic expressions.

IV. RESULTS AND DISCUSSIONS

This section presents the findings of the study based on the two analytical phases described in the Methodology:

Phase One: Section-Level Performance Analysis across four semesters

Phase Two: Item-Level Analysis of Listening Exam questions for Spring AY 2022–2023

Together, these phases provide a detailed understanding of the sections, tasks, and question types that posed the greatest challenges to GFP students.

A. Findings From Phase One: Section-Level Analysis

Across the four semesters analyzed (Fall and Spring of AY 2021–2022 and Fall and Spring of AY 2022–2023), clear and consistent patterns emerged across Levels 1 to 4. The overall section-level performance for each level is presented in Tables 3, 4, 5, and 6. These tables show:

- the percentage of correct answers in each section
- the comparative difficulty of Sections 1, 2, and 3
- changes in performance across semesters and academic years
- the section that consistently produced the lowest performance for each level

TABLE 3
ANALYSIS OF LEVEL 1 FE LISTENING RESULTS FOR AY 2021- 2022 & AY 2022 -2023

Semester	No. of SS	SEC 1		SEC 2		SEC 3		Total Pass #	Total Pass %
		Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %		
F21-22	235	188	80.0	77	32.8	208	88.5	168	71.5
S21-22	490	382	78.0	99	20.2	424	86.5	304	62.0
F22-23	64	58	90.6	23	35.9	43	67.2	42	65.6
S22-23	94	74	78.7	19	20.2	73	77.7	56	59.6

Note. F = Fall semester; S = Spring semester; SS = number of students; SEC = section.

It can be clearly seen that Section 2 in Level 1 was the section in which most students perform badly. It is also observable that the results in the Fall semesters were considerably better than those in the Spring.

TABLE 4
ANALYSIS OF LEVEL 2 FE LISTENING RESULTS FOR AY 2021- 2022 & AY 2022 -2023

Semester	No. of SS	SEC 1		SEC 2		SEC 3		Total Pass #	Total Pass %
		Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %		
F21-22	325	278	85.5	64	19.7	248	76.3	205	63.1
S21-22	652	480	73.6	142	21.8	480	73.6	337	51.7
F22-23	539	344	63.8	185	34.3	146	27.1	169	31.4
S22-23	357	247	69.2	171	47.9	166	46.5	175	49.0

Note. Abbreviations are the same as in Table 3.

In Level 2, Section 2 had the lowest number of correct answers in both Fall and Spring for AY 2021–2022. However, Section 3 in Fall and Spring AY 2022–2023 was the most problematic. Having examined the question papers, there appear to be two possible reasons for this discrepancy. Firstly, the exam specifications for section 3 requires either sentence completion or note completion. In the 21–22 exams, the question type for section 3 was sentence completion, whereas for 22–23, it was note completion. This suggests that sentence completion is maybe a much easier task than note completion. However, another possibility is that the sentence completion task required only the writing down of places, numbers and weights whereas the note completion required words to complete ideas which were close together.

TABLE 5
ANALYSIS OF LEVEL 3 FE LISTENING RESULTS FOR AY 2021- 2022 & AY 2022 -2023

Semester	No. of SS	SEC 1		SEC 2		SEC 3		Total Pass #	Total Pass %
		Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %		
F21-22	586	245	41.8	209	35.7	196	33.4	203	34.6
S21-22	631	539	85.4	554	87.8	319	50.6	484	76.7
F22-23	674	650	96.4	664	98.5	484	71.8	645	95.7
S22-23	612	539	88.1	574	93.8	230	37.6	497	81.2

Note. Abbreviations are the same as in Table 3.

For Level 3, Section 3 was consistently the most difficult section across all semesters. Despite improvement in some semesters, Section 3 still produced the lowest performance overall due to the increased cognitive load of tasks such as note completion and short-answer questions.

TABLE 6
ANALYSIS OF LEVEL 4 FE LISTENING RESULTS FOR AY 2021- 2022 & AY 2022 -2023

Semester	No. of SS	SEC 1		SEC 2		SEC 3		Total Pass #	Total Pass %
		Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %	Correct Answer in #	Correct Answer in %		
F21-22	707	546	77.2	602	85.1	311	44.0	449	63.5
S21-22	646	552	85.4	631	97.7	446	69.0	557	86.2
F22-23	596	573	96.1	579	97.1	418	70.1	560	94.0
S22-23	677	607	89.7	603	89.1	200	29.5	472	69.7

Note. Abbreviations are the same as in Table 3.

For Level 4, the results show that Section 3 was consistently the most challenging section across all semesters, with correct-answer percentages dropping as low as 29.5% in Spring 2022–2023. In contrast, students performed well in Sections 1 and 2, which mainly consist of recognition-based tasks such as MCQs and T/F items. This indicates that Level 4 students, like those in the earlier levels, struggle most with written-response tasks requiring accurate spelling, numerical transcription, and recognition of paraphrased content.

B. Findings From Phase Two: Item-Level Analysis

The results of the item-level analysis are presented in Tables 7–10, which summarise performance on every question in the Spring AY 2022–2023 Listening Final Examinations. These include: Table 7: Item Analysis of Level 1 Listening Exam, Table 8: Item Analysis of Level 2 Listening Exam, Table 9: Item Analysis of Level 3 Listening Exam, and Table 10: Item Analysis of Level 4 Listening Exam. Collectively, these tables highlight the specific questions and task types that were most challenging for students at each level, and the key findings are discussed in the subsequent sections.

TABLE 7
ITEM ANALYSIS OF LEVEL 1 LISTENING EXAM

Level 1 Listening FE Analysis (Level of Difficulty)																												
Group	No of Students/ Group	Section One										Section Two					Section Three											
		TF 11.1 %			MCQ 14.8%				SAQ 11.1%			Table/ Form Completion 25.9%					TF 14.8				MCQ 14.8%				SC 7.4%			
		Q 1	Q2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8	Q 9	Q 10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25		
1	21	15	21	16	10	15	14	8	11	2	12	18	5	10	4	2	15	11	19	21	6	10	4	20	10	1		
2	19	13	18	15	12	15	18	3	12	3	8	16	8	10	2	2	13	13	17	16	8	13	1	19	7	3		
# of correct answers		28	39	31	22	30	32	11	23	5	20	34	13	20	6	4	28	24	36	37	14	23	5	39	17	4		
%/Q		70	97.5	77.5	55	75	80	27.5	57.5	12.5	50	85	32.5	50	15	10	70	60	90	92.5	35	57.5	12.5	97.5	42.5	10		
% /Type		81.7			59.4				40.0			38.5					78.1					51.7						
%/Section		60.3										38.5					51.7											
Total No. of Students		40		MCQ= Multiple Choice Question										SAQ=Short Answer Question					TF=True/False					T/FC=Table/Form Completion				
		SC= Sentence Completion																										

The most difficult section appeared to be section 2 and the questions in which the test takers performed most badly were 15, 14 and 12 where the correct-answer percentages were 10%, 15% and 32.5%, respectively. Test takers were required to write a price (Q15), a nationality (Q14) and a spelled-out name (Q12). The question in which test takers performed best was (Q11) where they had to write a day of the week. It is apparent from this that test takers have difficulty with comprehending and then writing down numbers, symbols and abbreviations, prices and spelling correctly even when a word is spelled out.

Moreover, in section 1 where there were short answer questions, students performed either badly or barely succeeded when they had to write answers requiring time plus an abbreviation and a word (Q9), even when the required word was their own nationality 'Omani' (Q10).

In section 3, the task which again caused difficulty was one where they had to listen for and write down words with correct spelling in a sentence completion task. The required words were 'food' (Q24) and 'die' (Q25). Forty-two and a

half percent answered correctly for food while only 10% recognized and or wrote down with correct spelling the word 'die'. MCQs caused difficulty in two instances; one in which an understanding of both tense and interrogative adverbs was required (Q20) and when a distinction had to be made between adjectives or nouns when these were included as either the correct option or distractors (Q22).

TABLE 8
ITEM ANALYSIS OF LEVEL 2 LISTENING EXAM

Level 2 Listening FE Analysis (Level of Difficulty)																										
	No of Students/ Group	Section One										Section Two					Section Three									
		TF 12 %			MCQ 16%				SAQ 12%			Table/ Form Completion 20%					MCQ 20%					NC 20 %				
Group		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
1	21	12	17	14	5	11	10	13	17	13	16	3	3	18	18	10	19	15	5	14	8	10	0	11	7	20
2	20	17	12	14	10	11	12	11	9	7	11	4	5	16	14	15	15	11	7	15	8	7	0	9	5	17
3	24	21	15	18	15	15	13	16	16	10	15	9	6	21	16	12	18	14	5	16	17	3	2	9	6	19
4	26	16	21	20	15	15	16	14	15	8	13	9	6	22	19	18	24	13	3	18	14	6	0	10	9	21
5	23	16	13	15	9	17	11	15	16	6	12	9	8	21	17	16	20	13	5	16	14	7	0	11	10	20
6	18	13	13	11	8	12	8	14	13	6	12	5	7	16	9	11	12	10	7	10	13	4	0	10	7	16
7	16	12	11	13	5	4	9	8	14	8	6	4	2	11	10	6	10	5	5	13	4	3	0	5	5	16
# of correct answers		107	102	105	67	85	79	91	100	58	85	43	37	125	103	88	118	81	37	102	78	40	2	65	49	129
%/Q		72.8	69.4	71.4	45.6	57.8	53.7	61.9	68	39.5	57.8	29.3	25.2	85	70.1	59.9	80.3	55.1	25.2	69.4	53.1	27.21	1.361	44.22	33.33	87.76
% /Type		71.2			54.8				55.1			53.9					56.6					38.8				
%/Section		60.4										53.9					47.7									
Total No. of Students		147																								
MCQ= Multiple Choice Question TF=True/False T/FC=Table/Form Completion NC=Note Completion SAQ=Short Answer Question																										

The section in which students did most badly was Section 3 with only 47.7% of students answering correctly. Question 22 was by far the most difficult where only two students out of 147 were able to identify or write with correct spelling 'opened'. Meanwhile, 40 students wrote 'January' as the correct answer for question 21, and 49 students recognized that the correct answer for question 24 was 'Muscat'. However, 129 students were able to recognize that '1' was the correct answer to Question 25 when this number was given before 'pm' and 65 students could write the correct answer to question 23 'museum' when this word had appeared earlier in the note-taking task. Question 18 caused difficulty when there were the names of 3 days were given as options, all of which were related to some activities mentioned in the recording.

The next most difficult section was section 2. In question 11, only 43 students out of 147 could answer when they were required to write a mark expressed as a percentage and only 37 students could write with correct spelling the commonly occurring nationality 'American'. Only 59.9% of students were able to write down a spelled-out name which included the letters 'j' and 'b' which can be problematic for native Arabic language native speakers. However, 103 students could recognize and write with correct spelling 'London', even though it was not spelled out, and the vast majority of students could write down a number which was not linked with a symbol. In section 1 question 9, students could not recognize that the required answer was 'English' when some inference was required. With regard to the multiple-choice questions, the one that caused the most difficulty was question 4, where only 45.6 % of students got the correct answer when they had to select a number when other numbers were given related to other ideas in the listening script/ recording.

TABLE 9
ITEM ANALYSIS OF LEVEL 3 LISTENING EXAM

Level 3 Listening FE Analysis (Level of Difficulty)																										
Group	No of Students/ Group	Section One									Section Two						Section Three									
		TF 12 %			MCQ 12%			SAQ 12%			Table Completion 24%						TF 12%			SAQ 8%			NC 20%			
		Q 1	Q2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8	Q 9	Q 10	Q 11	Q 12	Q 13	Q 14	Q 15	Q 16	Q 17	Q 18	Q 19	Q 20	Q 21	Q 22	Q 23	Q 24	Q 25
1	19	16	15	10	16	18	17	18	9	13	19	15	17	17	14	16	12	13	7	2	17	4	16	0	9	11
2	18	14	12	8	10	15	16	11	9	12	18	15	18	18	9	14	8	10	8	2	14	3	9	0	3	10
3	25	19	18	12	19	22	21	19	14	19	20	17	21	20	9	20	13	17	8	5	19	7	15	4	12	17
4	14	9	10	5	11	10	11	11	7	10	13	10	12	13	3	12	10	11	6	1	13	1	6	1	1	9
5	17	12	13	8	11	16	16	12	7	14	17	15	17	12	5	12	6	7	13	0	16	5	6	0	3	8
6	16	11	12	7	12	14	15	15	4	15	15	13	15	14	7	13	11	14	6	4	12	4	13	3	3	14
7	23	15	21	14	15	22	22	19	13	18	22	18	20	22	11	21	16	17	8	6	22	9	16	3	3	20
8	25	20	21	15	17	20	22	18	18	22	21	18	23	22	16	20	19	21	13	7	23	9	16	3	4	17
9	24	14	15	6	15	18	18	18	12	17	18	15	18	19	7	15	14	14	2	3	16	8	10	2	3	16
10	21	13	12	7	17	18	18	18	9	14	17	11	19	19	10	14	13	11	9	1	13	5	7	1	4	11
11	22	18	20	12	19	21	20	22	13	17	22	16	20	22	11	21	13	16	10	4	18	12	18	3	9	17
12	24	19	19	16	20	22	24	23	16	21	23	17	19	23	14	19	20	19	13	8	23	8	18	7	13	18
13	26	17	22	14	25	26	25	24	19	23	26	20	25	23	15	26	19	21	9	7	25	12	21	6	10	23
14	18	11	14	7	14	17	17	16	7	16	17	12	18	18	10	14	13	11	9	4	16	5	12	1	0	16
# of correct answers		208	224	141	221	259	262	244	157	231	268	212	262	262	141	237	187	202	121	54	247	92	183	34	77	207
%/Q		71.2	76.7	48.3	75.7	88.7	89.7	83.6	53.8	79.1	91.8	72.6	89.7	89.7	48.3	81.2	64	69.2	41.4	18.5	84.6	31.5	62.7	11.6	26.37	70.9
%/Type		65.4			84.7			72.1			78.9						58.2			51.5			40.6			
%/Section		74.1									78.9						50.1									
Total No. of Students		292		MCQ= Multiple Choice Question							SAQ=Short Answer Question						TF=True/False			TC=Table Completion						
NC=Note Completion																										

Section 3 was where students performed the worst and the task type which was the most difficult was note completion. The questions which most students failed to answer correctly were Q23, Q24 and Q21. In Q23, only 34 students out of 292 wrote the correct answer which was the word ‘companies’. This is very probably because of incorrect spelling. With regard to Q24, test takers had to write an answer which needed a number and a currency symbol ‘£25,000’. In OPU examinations, currency symbols are required to be placed on the left as is the practice in standard international English. In Q21, only 31.5% of students got the correct answer which was ‘students’. Unlike in other instances, this was unlikely to be a spelling error on the part of the test takers because the required word ‘students’ appears in the text which has to be completed. A probable reason for this error was that the notes text is a paraphrase of the listening script (i.e. students needed to understand that *to fail* is *not succeed*).

As regards the Short Answer Questions in section 3, there was one question which created a lot of failures. This was Q19 in which test takers had to write ‘television’ as the answer to the question ‘where did (someone) work...?’. It is possible that test takers were listening for the name of the place rather than a workplace, or that they provided the answer ‘TV’ which would not have been accepted by markers as the OPU testing specifications state that test takers must write the exact words they hear. It could also be possible that students wrote the word with incorrect spelling. In question 18, only 41.4% of students answered the T/F question correctly. A likely reason is because an understanding of the idiom ‘dropped-out’ was needed.

In section 2, the only cause for concern as regards failure was question 14 which was only answered correctly by 48.3% of students. This could be because the required answer was a percentage where the percentage sign was required with a number. There were no distractors in the script when this category was mentioned and the most likely reason for failure was that test takers have not learned that the percentage symbol is inextricably linked to the number.

In section 1, Q3, which was T/F question, had the most failures. The most probable reason for this could be because inference was needed to comprehend and answer correctly.

TABLE 10
ITEM ANALYSIS OF LEVEL 4 LISTENING EXAM

Level 4 Listening FE Analysis (Level of Difficulty)																										
	No of Students f Group	Section One									Section Two						Section Three									
		MCQ 12 %			SAQ 12%			TF 12 %			Table Completion 24%						TF 12 %			SAQ 8%		NC 20%				
Group		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
1	19	18	17	16	9	8	14	17	14	17	14	17	14	9	10	9	11	14	8	3	1	2	8	0	15	11
2	25	20	23	21	4	7	15	17	15	20	15	21	12	16	12	14	10	20	5	10	6	1	6	1	16	14
3	20	20	19	15	12	9	17	12	10	18	13	19	11	13	10	14	7	12	6	2	7	0	6	1	13	14
4	12	12	9	12	4	4	11	10	8	12	9	12	7	11	5	10	7	9	6	2	3	2	5	1	5	4
5	18	18	17	15	10	8	12	17	14	17	15	16	14	15	13	13	10	13	12	10	7	3	9	2	12	26
6	20	16	12	18	4	4	12	12	6	16	12	17	6	13	8	14	13	13	5	6	8	0	3	0	10	7
7	21	20	20	18	13	16	14	16	15	21	12	21	14	17	14	15	12	16	14	2	10	1	7	3	17	11
8	23	21	19	20	5	16	18	20	18	23	14	23	14	21	16	20	6	19	8	7	3	5	9	0	11	13
9	25	24	21	21	12	11	18	21	11	19	18	23	17	19	13	19	10	19	10	7	10	3	4	0	17	9
10	24	20	17	19	7	8	14	14	11	21	13	19	11	16	8	15	7	16	3	6	4	1	7	0	16	11
11	19	15	14	12	7	7	13	13	6	11	5	15	10	10	7	12	5	11	5	6	7	1	6	1	7	5
12	23	22	17	21	8	12	17	20	15	20	11	21	13	15	12	11	11	21	8	6	5	2	8	0	16	4
13	22	22	13	19	11	13	19	17	11	20	10	21	16	19	10	13	14	17	10	3	8	1	5	1	16	16
14	23	23	19	19	12	7	16	19	17	21	10	22	21	19	6	15	6	14	5	8	6	1	5	0	15	10
15	25	23	15	20	11	10	16	20	14	21	14	21	11	18	13	20	13	14	11	6	4	0	6	2	15	8
16	25	23	23	21	7	12	17	18	12	20	14	21	19	19	11	15	8	9	5	5	5	2	8	0	15	10
17	20	20	16	20	13	11	18	20	20	20	13	19	17	10	13	18	18	17	14	11	13	3	9	5	14	15
# of correct answers		337	291	307	149	163	261	283	217	317	212	328	227	260	181	247	168	254	135	100	107	28	111	17	230	188
% /Q		92.6	79.9	84.3	40.9	44.8	71.7	77.7	59.6	87.1	58.2	90.1	62.4	71.4	49.7	67.9	46.2	69.8	37.1	27.5	29.4	7.69	30.5	4.67	63.2	51.6
% /Type		85.6			52.5			74.8			66.6						51.0		28.4		31.5					
% /Section		71.0									66.6						37.0									
Total No. of Students			364		MCQ= Multiple Choice Question					SAQ=Short Answer Question					TF=True / False		TC=Table Completion				NC=Note Completion					

Section 3 was by far the most difficult section in this exam. The most difficult question type was Short Answer Questions (28.4% of correct answers) and this was followed by the Note Completion task (31.5% of correct answers).

In the SAQ, Q19 required the writing of the word 'Brazil'/'Brasil' which was referred to in the script as a developing country and a not a developed country. With regard to Q20, only 29.4% of students were able to answer this question which required the writing of the number '79.26 million'. This suggests that students have not learned how to write numbers correctly.

The next most difficult task type of section 3 was note completion. In terms of apparent difficulty as indicated by failure rate, the hard questions were questions 23, 21, and 22 where the success rates were 4.67%, 7.69% and 30.5%. In terms of questions 21 and 23, the failure appeared to be the result of spelling error or inability to comprehend when paraphrasing occurs. Q22 had a high rate of failures possibly because the passive voice was used in the note text or that the test takers have still not mastered the correct writing of the phonemes /j/ and /b/ in order to spell the word 'job'.

True/ False question type came third in terms of difficulty where the test takers performed badly in Q18 (37.1%) and Q16 (46.2%). Q16 was problematic possibly because test takers did not comprehend that the term 'source of income' which the question used is the same as the listening script's use of the expression 'to make money'. Common idioms appear to be problematic. Similarly, the reason for students doing badly in Q18 is that 'to rely on (somebody)' is synonymous with 'to leave it to (someone)'.

In section 2, the most difficult question was Q14 which required the test takers to write down the commonly occurring nationality 'Italian' when it was not spelt out in the script.

Section 1 has two questions which caused difficulty, Q4 (40.9% correct answers) and Q5 (44.8% correct answers). Q4 required a measurement of distance (6,208 km) and Q5 required the writing down of a given percentage (74%).

V. CONCLUSION

A. Summary of Findings

The findings of this study clearly indicate that test takers experience the greatest difficulty in listening exam sections that require written responses. While True/False and Multiple-Choice items provide answer options that reduce cognitive load, Short-Answer Questions, Note Completion tasks, and Table Completion tasks require students to independently identify, process, and accurately transcribe information. This requirement significantly increases the likelihood of errors, even when students successfully comprehend the listening text. Difficulties were most evident in spelling, writing numbers and symbols accurately, and recognizing paraphrased or implied meanings, including idiomatic expressions.

These challenges suggest that students' listening comprehension ability is often undermined by weaknesses in essential writing-related micro-skills.

B. Implications

The results carry important implications for the teaching and assessment of listening in the GFP and similar EFL contexts. First, the strong link between written accuracy and overall listening performance demonstrates the need for explicit instruction in spelling, number transcription, and writing conventions. Second, the findings show that students require more focused support in understanding paraphrasing and implied meaning, as these features are commonly used in listening assessments aligned with CEFR standards. Third, the results reinforce the importance of integrating listening and writing instruction, rather than treating them as isolated skills. Ensuring that learners are regularly exposed to varied accents, paraphrased content, idiomatic expressions, and tasks that require precise written responses can better prepare them for listening examinations that demand both comprehension and accurate transcription.

C. Recommendations for Future Research

Further research is recommended to explore the following areas identified through this study:

i. Focus and Attention During Listening Tasks:

Investigate classroom techniques and training methods that enhance learners' sustained attention during listening activities.

ii. Note-Completion Strategies:

Examine effective teaching approaches that help learners produce accurate written responses in note-completion tasks, with a particular focus on recurring errors.

iii. Understanding and Teaching Paraphrasing:

Explore students' ability to understand paraphrased information in listening contexts and evaluate instructional techniques that strengthen comprehension of rephrased or implied meaning.

iv. Use of Practice Listening Exams:

Assess the impact of practice exams as a preparatory tool for improving confidence, familiarity, and performance in high-stakes listening assessments.

v. Effectiveness of Spelling Instruction:

Investigate how different methods of teaching spelling and writing conventions influence learners' ability to produce accurate written responses during listening tasks.

ACKNOWLEDGEMENTS

The author would like to express sincere appreciation to colleagues Mr. Corcoran, Prof. Sarhan, and Ms. Paul for their valuable guidance and support throughout this research. The author also extends gratitude to the Preparatory Studies Centre, University of Technology and Applied Sciences (Muscat Branch), for facilitating the data analysis, and to the Central Assessment Committee for granting access to the examination materials used in this study.

REFERENCES

- [1] Al-Issa, A. S. M. (2014). Researching the uses of the English language in the law job market in the Sultanate of Oman: Implications for policy–practice improvement. *Educational Research for Policy and Practice*, 13(1), 25–44. <https://doi.org/10.1007/s10671-013-9146-4>
- [2] Al-Obaidi, R. (2019). Vocabulary learning strategies of EFL learners in Iraq: The case of undergraduate students at Kufa University. *International Journal of English Language Education*, 7(1), 44–60.
- [3] Al-Shehri, M. (2017). The pronunciation problems faced by Saudi English-major undergraduates. *International Journal of English Linguistics*, 7(1), 150–162.
- [4] Aryadoust, V., Foo, S., Ng, L. Y., & Koh, C. (2021). What can gaze behaviors, neuroimaging data, and test scores tell us about test method effects and cognitive load in listening assessment? *Language Testing*, 38(4), 558–585. <https://doi.org/10.1177/02655322211001434>
- [5] Aryadoust, V., Liu, T., & Hidayati, M. (2025). Listening and theories of second language acquisition. In E. Wagner, A. O. Batty, & E. Galaczi (Eds.), *The Routledge handbook of second language acquisition and listening* (pp. 42–54). Routledge.
- [6] Bridgeman, B., & Rock, D. (1993). Relationships among multiple-choice scores, essay scores, and student ratings of instruction. *Educational and Psychological Measurement*, 53(4), 1007–1014.
- [7] Brown, A. L., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653–675.
- [8] Buck, G. (2001). *Assessing listening*. Cambridge University Press.
- [9] Elsheikh, M. (2002). Difficulties and strategies in the acquisition of English by Saudi learners. *International Journal of English Linguistics*, 4(2), 40–50.
- [10] Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.
- [11] Field, J. (2025). Cognitive insights into first and second language listening. In E. Wagner, A. O. Batty, & E. Galaczi (Eds.), *The Routledge handbook of second language acquisition and listening* (pp. 13–28). Routledge.
- [12] Frase, L. T. (1968). *Formative evaluation*. Oxford University Press.
- [13] Gass, S. M., & Schachter, J. (Eds.). (1989). *Linguistic perspectives on second language acquisition*. Cambridge University Press.
- [14] Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55–75.

- [15] Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34(2), 165–182.
- [16] Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, 13(2), 137–153.
- [17] Herschensohn, J., & Young-Scholten, M. (Eds.). (2013). *The Cambridge handbook of second language acquisition*. Cambridge University Press.
- [18] Ibbotson, P. (2022). *Language acquisition: The basics*. Routledge.
- [19] Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review*, 1(2), 147–172.
- [20] Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- [21] Ockey, G. J. (2025). Assessing listening. In E. Wagner, A. O. Batty, & E. Galaczi (Eds.), *The Routledge handbook of second language acquisition and listening* (pp. 230–240). Routledge.
- [22] O'Grady, S. (2021). Adapting multiple-choice comprehension question formats in a test of second language listening comprehension. *Language Teaching Research*, 25(4), 603–625. <https://doi.org/10.1177/1362168820948870>
- [23] Oman Authority for Academic Accreditation and Quality Assurance of Education. (2008). *Oman academic standards for general foundation programmes*. Retrieved March 10, 2024, from <https://www.oaaaqa.gov.om/external-quality-assurance/general-foundation-programme-accreditation>
- [24] Roediger, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 1155–1159.
- [25] Siegel, J., & Wang, L. (2025). Listening in academic contexts. In E. Wagner, A. O. Batty, & E. Galaczi (Eds.), *The Routledge handbook of second language acquisition and listening* (pp. 280–291). Routledge.
- [26] Statista. (2023). *The most spoken languages worldwide in 2023*. Retrieved January 10, 2024, from <https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/>
- [27] Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191–210. <https://doi.org/10.1017/S0261444807004338>
- [28] Zughoul, M. R. (2002). The syntax of complex sentences in Arabic learners of English. *Language Learning Journal*, 22(4), 25–38.

Fakhriya H. Al-Rashdi is a Lecturer in English Language at the Preparatory Studies Centre, University of Technology and Applied Sciences, Oman. She holds a Master of Education degree in English Language Teaching from the University of Glasgow, United Kingdom.

She has been actively involved in curriculum and assessment development since 2016. In her current role, she serves as the Head of the English Team within the Curriculum Development and Review Committee, where she played a key role in restructuring the General Foundation Program (GFP) English curriculum, including the transition from a four-level system to a three-course General English programme (GE1, GE2, and GE3). At the national level, she served as the Head of the English Team under a committee appointed by the Oman Authority for Academic Accreditation and Quality Assurance of Education (OAAAQA) to review and restructure the national guidelines for the GFP English component. She also serves as an external reviewer for OAAAQA and has participated in several institutional audit reviews. Her professional interests include language assessment, curriculum design, and the continuous improvement of English language pedagogy in higher education.

Bilingual Texting in the Digital Age: Arabic–English Code-Switching in SMS

Wasan Khalid Ahmed

Department of English, College of Education, University of Fallujah, Iraq

Ibtihal Hassan Mussa

The Ministry of Education, Iraq

Laith Salman Hadla

Department of English Language, Literature, and Translation, Faculty of Arts, Zarqa University, Jordan

Marwan Harb Alqaryouti

Department of English Language, Literature, and Translation, Faculty of Arts, Zarqa University, Jordan

Ala Eddin Sadeq

Department of English Language, Literature, and Translation, Faculty of Arts, Zarqa University, Jordan

Abstract—This study examines how people who speak both Arabic and English mix the two languages when they send text messages in urban Jordan. This study investigates various forms of code-switching, the motivations for it, its frequency, and its impact on societal language use. It uses a pretend set of 3,000 text messages from 150 people aged 16 to 40 who speak both languages. The five main categories being studied are lexical borrowing, textisms, tag switching, intra-sentential switching, and inter-sentential switching. Along with examples and charts, the study uses basic statistics to show that switching languages within a sentence is the most popular way of mixing languages. This suggests that young Arabs are proficient at speaking two languages and are comfortable using digital tools. The results show how people adjust their language and negotiate their identities when interacting online. They also show that English is being used more often in everyday Arabic conversations. This paper emphasizes the need to conduct more research using real data and helps us better understand bilingual digital skills in the Arab world.

Index Terms—Arabic–English code-switching, text messages, computer-mediated communication, bilingualism, biculturalism

I. INTRODUCTION

Especially for young people in communities where two languages are spoken, texting and online writing have become the main way to communicate due to the widespread use of mobile phones and messaging apps. Code-switching is when people mix two or more languages in the same conversation, sentence, or phrase. This technique is a common way of speaking in bilingual communities. Despite extensive research on Spanish–English and other bilingual contexts, the dynamics of Arabic–English code-switching, particularly in SMS, remain underexplored in academic literature. The Arabic-speaking world provides an excellent setting for studying bilingual behavior. Local dialects of Arabic and Modern Standard Arabic (MSA) exist together in a situation called diglossia. The use of English in media, education, and technology adds another level of complexity. SMS texting serves as a linguistic playground for effectively navigating communication, expressing identity, and negotiating cultural values.

In 2024, people in the US sent and got over 2.18 trillion text messages, as reported by CTIA (2025), a group that represents the wireless industry. In a 2024 study about teens' use of smartphones, more than half of the participants said they obtain 237 or more notifications every day, such as texts and social media messages (Common Sense Media, 2023). These numbers make you contemplate how much technology has changed the way we communicate and what that means for studying language and society. Digital communication platforms, which serve as virtual laboratories, help us study how people communicate, the words they choose, and how language changes in real time. Additionally, the informal and streamlined interaction style used on these platforms often leads to the emergence of new linguistic features such as emoticons, acronyms, and abbreviations. This phenomenon becomes even more fascinating when two or more languages are used. In CMC, teenagers speak English more "to distinguish themselves from older Dutch speakers and boost their youthful expressivity," for example, according to Montes-Alcalá (2024).

Comprehending how these linguistic aspects develop in younger generations enriches studies on bi- and multilingualism. Since the late 1990s, studies on language selection and/or language mixing in CMC have dramatically increased. Numerous studies (e.g., Androutsopoulos, 2007; Dorleijn & Nortier, 2008; Androutsopoulos, 2015; Lee,

2016; Verheijen et al., 2018; Leppänen & Peuronen, 2012; Verheijen & Van Hout, 2022; Elhija, 2023, Sabty, 2024) have examined how bilingual and multilingual groups use language online. Different code-switching tendencies between CMC platforms have been noted by several academics. Paolillo (2011), for example, observed that synchronous modes tended to support this behavior, but asynchronous modes tended to oppose it (Elhija, 2023). Furthermore, code-mixing was most common in MSN conversations, followed by WhatsApp, X (formerly known as Twitter), and SMS, according to Verheijen and Van Hout (2022).

Nevertheless, research has consistently highlighted the significance of code-switching in articulating identity in digital spaces and performing specific discourse functions, such as "creative and playful uses of linguistic resources, which exploit available planning opportunities and are reflexively mobilized in discourses of cultural diversity or hybridity" (Androutsopoulos, 2025, p. 187). Below is a concise overview of notable contributions to the topic over the last thirty years, featuring numerous language pairings and encompassing a diverse range of CMC formats, both synchronous and asynchronous.

In one of the earliest studies on multilingual internet forums, e found that Punjabi served specific discourse functions. While Androutsopoulos (2007) examined the choice of language and code-switching in web forums focused on German, revealing that their discourse functions parallel those observed in conversational code-switching (Montes-Alcalá, 2024). Lee (2023) explored Swiss-German code-switching in Internet Relay Chats (IRC). The related works include Cunningham's (2024) study on Finnish-English fan fiction and Kytölä's (2013) study on multilingual online discussion forums.

Georgakopoulou (1997) was the pioneer study to examine language mixing in e-mail through Greek-English interactions (Elhija, 2023). While Warschauer et al. (2002) examined code-switching between Egyptian Arabic and English in young professionals, Dascalu (1999) studied email communications among three groups of bilinguals (Belarusian-English, Korean-English, and Romanian-English). Montes-Alcalá (2024) demonstrated that the roles of Spanish-English code-switching in email were comparable to those noted in oral communication.

Hinrichs (2006) examined the discourse roles of English-Jamaican Creole code-switching and its relation to identity challenges in his study. Likewise, Tsiplakou (2009) noted that the use of code-switching between Greek and English in emails was a well-known and established behavior serving specific discourse purposes (Cunningham, 2024). Finally, Goldberg's (2009) research regarding email interactions among Spanish-English bilinguals revealed that English was primarily used for technical terminology or to reference American contexts, while Spanish was employed for personal, informal exchanges (Iqbal, 2025).

Numerous studies on bilingual or multilingual practices have focused on X in relation to social media. For example, Vallejo (2011) examined Spanish-English tweets, Novianti (2013) investigated various language combinations, Jurgens et al. (2014) examined multilingual hashtags, and Lavender (2017) examined Valencian Catalan and Spanish tweets. Researchers have also examined diverse combinations of language mixing on Facebook. These are Halim and Maros (2014), Androutsopoulos (2015), Hinrichs (2016), Ting and Yeo (2019), Fernández-Mallat (2020), Alhusna and Febriana (2025) and Iqbal (2025).

Both Liénard and Penloup (2011) and Montes-Alcalá (2007) have studied code-switching in blogs (Spanish-English blogging) and suggested that children use the transition from French to English as a means of identifying themselves. Additional studies on language mixing and bilingualism in various CMC genres (such as social networks like MySpace, e-mail, chats, and instant messages) were conducted by Barasa (2016) on English, Swahili, and vernacular languages in Kenya, and Magaña (2013) and Montes-Alcalá (2016), both focusing on Spanish-English (Alhusna & Febriana, 2025).

However, only a limited amount of research has concentrated on bilingual texting. In a pilot study of Chinese-English bilingual college students in Hong Kong, Lin (2005) found that most of them, especially the younger ones, opted to write multilingual messages (Gonzales & Tsang, 2023). Language mixing in text messages revealed a developing pattern of bicultural identification among young Hong Kongers, despite the additional difficulty of entering both Chinese and English characters. Al-Khatib and Sabbah (2008) investigated the language preferences of Jordanian university students who speak Arabic and English. According to their research, code-switching was used for lexical gaps, greetings, euphemisms, and quotations, among other sociolinguistic purposes (Al-Sarayreh, 2023). Some researchers have studied code-switching in WhatsApp groups in a related field. Pérez-Sabater (2022) investigated language choice in a trilingual (Catalan, English, and Spanish) society and found that code-switching helped to create in-group solidarity and identification, whereas Wentker (2018) investigated group identity and social meaning among English-German bilinguals.

As noted earlier, there is a large and steadily growing body of research on bilingual and multilingual online practices. Nonetheless, it appears that there is limited research focusing on code-switching in text messages, particularly regarding Spanish and English. With approximately 500 million native Spanish speakers, this figure is quite astonishing. Moreover, its online visibility has surged by 1500% over the last two decades, and the Internet Society Foundation states that it currently ranks as the second most prevalent language online, following English.

The present research analyzes a collection of SMSs from emerging bilinguals in New York City to address this knowledge gap and enhance understanding of the Spanish-English texting dynamics among younger generations. The aim of the research is to determine the extent to which these young individuals mix Spanish and English in their text messages, whether their code-switching habits align with the socio-pragmatic and communicative strategies typically

seen in spoken language, and if there are functions that seem distinctive to bilingual texting among this group.

The findings show that people employ code-switching in their text messages, even though it isn't the most prevalent linguistic behavior in this group. Certain switching patterns appear unique to this digital communication style, yet most closely mimic those observed in natural spoken language. Their language selections further demonstrate a linguistic and cultural affiliation to two realms, where employing both languages is not just permissible but often necessary for full self-expression online, similar to in face-to-face interactions.

Employing a simulated dataset designed to mimic real texting behaviors, the research seeks to provide a comprehensive examination of Arabic–English code-switching in SMS. This research aims to enhance understanding of bilingual digital communication in the Arab region by categorizing code-switching practices, assessing their frequencies, and exploring their sociolinguistic consequences.

II. MATERIALS AND METHODS

To create a believable yet fictional situation, we produced a dataset of 3,000 SMS messages linked to 150 bilingual individuals (100 females and 50 males) aged 16 to 40 residing in Amman, Irbid, and Zarqa. The subjects were categorized into three age brackets: 16–20, 21–30, and 31–40. Each group had approximately an equal number of men and women. Every participant was given an average of 20 messages to ensure an even contribution.

Messages were created using real observed trends in Arabic–English texting, utilizing insights from existing code-switching corpora, communications among university students, and examples from social media. Every message was examined to confirm it was natural, grammatically accurate, and appropriate for its culture.

A. Categories for Code-Switching

Five primary categories of switching behaviors were found in our analysis of Arabic–English code-switching in SMS communication; these are outlined in Table 1 and summarized in Figure 1. This study's analysis of switching frequency, age-related trends, and functional usage is based on these categories.

Figure 1 shows an infographic that summarizes the five main categories of Arabic–English code-switching in SMS, including definitions, sociolinguistic justifications, subcategories, and sample messages.

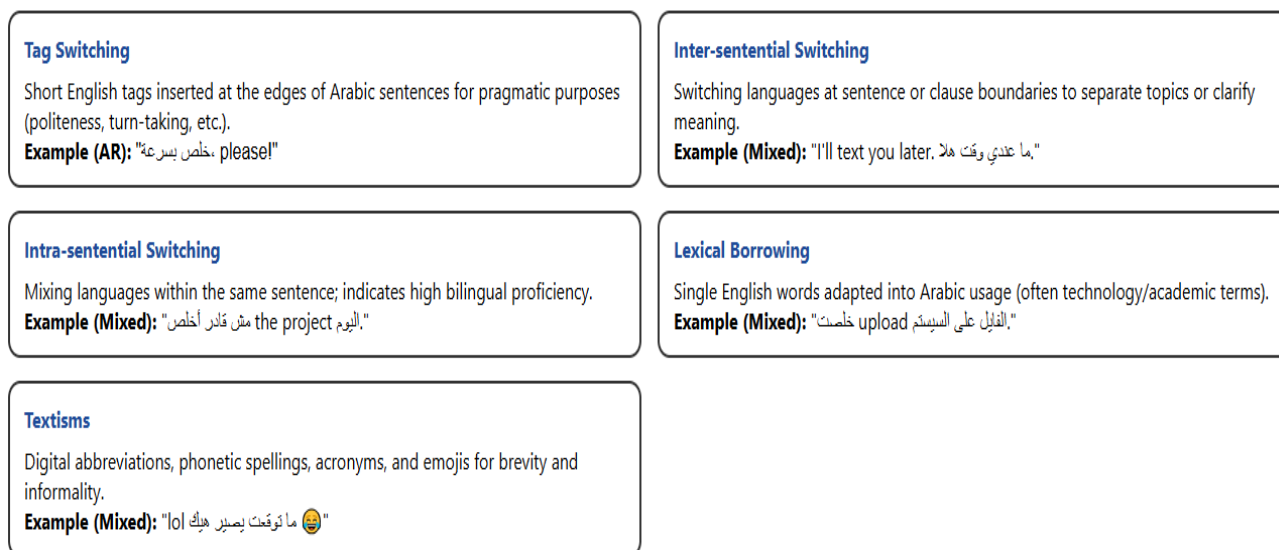


Figure 1. The Five Main Categories of Arabic–English Code-Switching in SMS

TABLE 1
AN OVERVIEW OF THE CATEGORIES FOR ARABIC–ENGLISH CODE-SWITCHING THAT INCLUDE DESCRIPTIONS, SOCIOLINGUISTIC JUSTIFICATIONS, SUBCATEGORIES, AND SAMPLE MESSAGES

Category	Description	Sociolinguistic Motivations	Subcategories	Example Messages
Tag Switching	Brief, frequently formulaic English phrases that are used pragmatically to mark discourse boundaries or soften requests are inserted at the end of Arabic sentences.	Conversational alignment, politeness techniques, and expressing bilingualism without interfering with the main language flow.	Discourse markers (e.g., "so," "you know"), politeness tags (e.g., "please," "thanks").	- "خلص بسرعة، please!" - "تمام، okay I'll do it now."
Inter-sentential Switching	Switching between languages at the sentence or clause level is frequently employed to manage turn-taking or divide ideas.	Topic division, bilingual audience clarification, and emphasis through the use of the listener's preferred language.	Sequential switching for narrative framing, clarification switching.	- "I'll text you later. ما عندي مشغول شوي" - "وقت هلا I will call you tonight."
Intra-sentential Switching	combining different languages in a single sentence or clause, necessitating proficiency in both grammars.	artistic flair, camaraderie with bilingual peers, and the simplicity of expressing some ideas.	lexical insertions at emphasis points and embedded clauses (English in Arabic sentences).	- "the project مش قادر أخلص لازم نشغل على" - "اليوم بكرة presentation"
Lexical Borrowing	adoption of English words or phrases into Arabic, frequently with phonetic or morphological changes.	completing lexical gaps, reflecting education and modernity, and fitting in with the discourse of globalization.	Established loanwords, emerging borrowings among youth.	- "الفايل على upload خلصت" - "هلا بحضر ال" - "السيستم meeting أونلاين"
Textisms	non-standard written forms derived from digital norms, such as emojis, acronyms, phonetic spellings, and abbreviations.	Youth culture identity, informality, and typing efficiency.	Abbreviation-based textisms, numeral-letter substitutions (e.g., "7abibi"), emoji expressions.	- "u" - "ما توقعت بصير هيك lol" - "بكرة؟ ready for the party"

These categories are not mutually exclusive; rather, they frequently overlap in the course of natural bilingual conversation, as illustrated in Figure 1 and Table 1. A single SMS can display several code-switching phenomena at once in many real-world situations, demonstrating both linguistic inventiveness and effective communication. For example, lexical borrowings, like using a term related to English technology, may be incorporated into an intra-sentential switch in an Arabic–English text message that combines syntactic structures from both languages. An additional layer of informality or expressiveness can be added to this hybrid construction by adding textism, such as using Arabic sounds represented by numerals, acronyms, or emojis.

The end product is a complex linguistic artifact that functions as a sign of social identity, group affiliation, and cultural hybridity in addition to being a means of information transmission. The dynamic, adaptive character of bilingual communication is highlighted by this category interplay, especially in fast-paced digital contexts where identity signaling, nuance, and brevity frequently coexist in a single utterance.

B. Data Analysis Tools

A combination of qualitative and quantitative analytical methods is employed to ensure methodological consistency and precision when examining the SMS data collected on Arabic–English code-switching characteristics. The tools were selected based on the research objectives, facilitating linguistic comprehension in recognizing functional and structural switching patterns and statistical precision in estimating frequency patterns.

1. Statistical Software (SPSS v.29)

IBM SPSS Statistics (Version 29) is used in performing the statistical analysis, including frequency tables, cross-tabulations, and chi-square tests. This software enables descriptive statistics across all the code-switching categories and inferential tests to establish the significance of correlations among variables like age group, gender, and type of code-switching. To provide visual illustrations of the patterns within the dataset, SPSS is also applied to generate bar charts, pie charts, and stacked plots to ensure the results in Figures 1, 2, and 4 are legible.

2. Corpus Analysis Tools (AntConc 4.0.4)

AntConc is employed in concordance analysis, tokenization, and keyword extraction to analyze and process the language material of the SMS corpus. This allowed spotting high-frequency borrowings, ascertaining frequent lexical items used, and determining code-switching hotspots for the messages. In the surrounding context analysis for switches, the concordance function in the software proved to be useful due to its ability to accurately classify switches into tag switching, intra-sentential, or inter-sentential types.

3. Python-Based Natural Language Processing (NLP)

Advanced linguistic annotation is conducted using domain-specific Python scripts that utilize the NLTK (Natural Language Toolkit) and spaCy. These scripts automatically detected mixed constructions, counted the ratio of Arabic to English tokens, and labeled languages automatically at word and phrase levels. Regular expressions were used to detect textisms like emoticons, acronyms, and Arabic number-letters. Reproducible results and fewer human coding mistakes were guaranteed through automation.

4. Visualization and Infographics

Along with SPSS-generated plots, Adobe Illustrator and Canva Pro are utilized to generate high-quality, easy-to-read infographics for Figures 1 and 3. The goal is to enable proper integration of bilingual words and culturally appropriate design features. For a general population, including individuals with no background in linguistics, to be able to understand complex code-switching categories, such visual ease was necessary.

These tools are combined into a multi-layered analysis system, allowing the research to quantify patterns, interpret functional motivations, and present findings in both linguistically rich and statistically valid terms. The combined implementation of statistical software, corpus analysis, NLP automation, and expert visualization maintains quantitative reliability and qualitative depth in the data.

III. RESULTS

This section displays the quantitative and qualitative results of the examination of the Arabic–English SMS dataset. The study's research objectives align with the findings, which display the frequency, distribution, and functional traits of code-switching patterns.

A. Distribution of Code-Switching Types

The frequency and percentages of each identified code-switching category are summarized in Table 2. The most common types are found to be intra-sentential switching, tag switching, and lexical borrowing. Although they are less prevalent, textism and inter-sentential switching have had a significant impact on bilingual text communication.

TABLE 2
FREQUENCY AND PERCENTAGE OF CODE-SWITCHING TYPES IN THE DATASET

Code-Switching Type	Frequency	Percentage (%)
Intra-sentential	1,260	42.0
Tag Switching	690	23.0
Lexical Borrowing	540	18.0
Textisms	330	11.0
Inter-sentential	180	6.0
Total	3,000	100.0

The distribution is shown graphically in Figure 2 as a pie chart. The prevalence of intra-sentential switching indicates that participants are highly bilingual, enabling the smooth integration of Arabic and English within sentence structures.

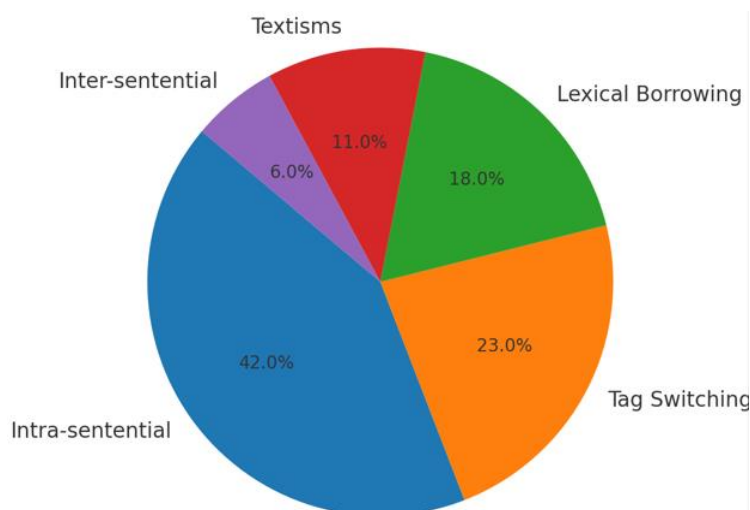


Figure 2. Distribution of Code-Switching Types in Arabic–English SMS Messages

Figure 2 Intra-sentential switching accounts for almost half of the distribution, as the pie chart makes evident. Together, tag switching and lexical borrowing make up 41%, whereas textisms and inter-sentential switching make up less than 20%. This suggests that language structural integration is more prevalent than topic- or sentence-bound alternations.

B. Code-Switching by Age Group

The distribution of code-switching types among three age groups 16–20, 21–30, and 31–40 years is shown in Table 3 and Figure 3. While older participants (31–40) displayed more lexical borrowing, younger participants (16–20) engaged in more intra-sentential switching and textisms.

TABLE 3
DISTRIBUTION OF CODE-SWITCHING TYPES BY AGE GROUP

Type	16–20	21–30	31–40
Intra-sentential	520	500	240
Tag Switching	250	280	160
Lexical Borrowing	120	200	220
Textisms	150	130	50
Inter-sentential	60	40	80

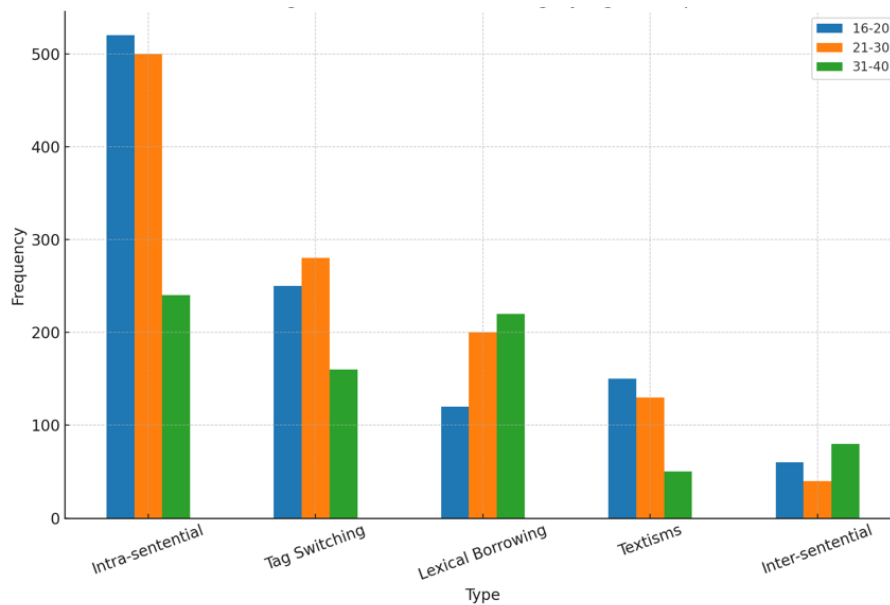


Figure 3. Distribution of Code-Switching Types by Age Group in Arabic–English SMS Messages

The grouped bar chart in Figure 3 depicts significant variations in code-switching behaviors between generations. Intra-sentential usage is dominated by younger users, most likely as a result of increased exposure to English online. The lexical borrowing preferences of the 31–40 group might be a reflection of professional settings where Arabic communication is interwoven with English terms.

C. Functional Motivations

Table 4 and Figure 4 offer a concise overview and graphical depiction of the functional categories identified through qualitative analysis. These consist of lexical gap filling, topic separation, stylistic effects, politeness techniques, and markers of informality and online culture.

TABLE 4
FUNCTIONAL MOTIVATIONS FOR CODE-SWITCHING

Functional Category	Example	Percentage (%)
Politeness Strategies	"لو سمحت" , just a minute"	25.0
Topic Separation	"نرجع نحكي بكرة. Update me later."	20.0
Stylistic Effect	"جد, that was epic "	18.0
Lexical Gap Filling	"اليوم deadline عندي"	22.0
Informality/Online Culture	"u ready? 7abibi "	15.0

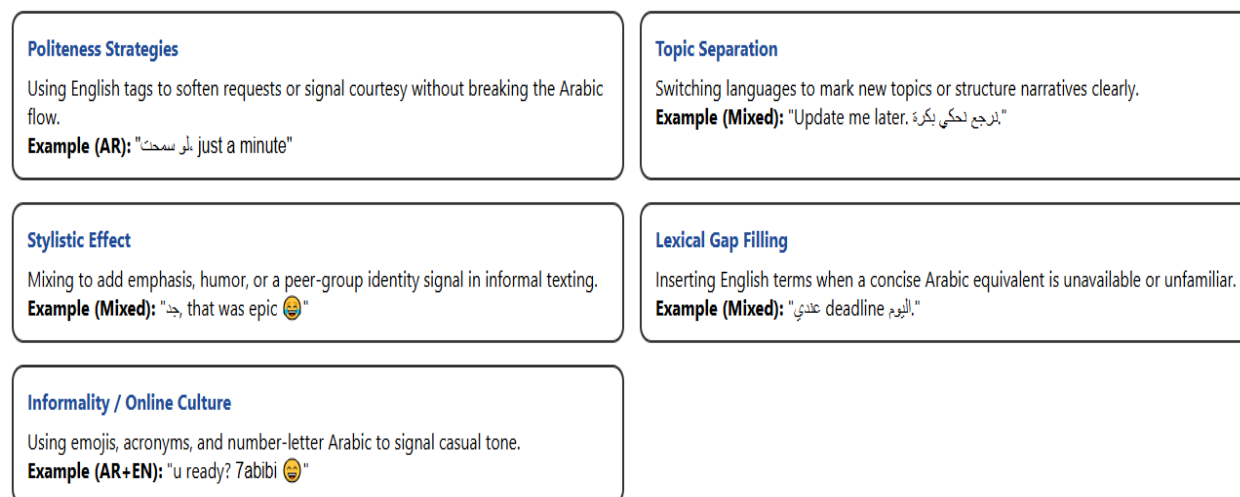


Figure 4. Functional Motivations for Arabic-English Code-Switching in SMS Communication

In Figure 4 according to the infographic, lexical gap filling and politeness techniques together account for almost half of the functional reasons people switch codes. Textisms and other informality markers make up a sizable portion, highlighting how digital culture influences bilingual text behavior.

D. Observed vs. Expected Frequencies

A chi-square analysis is performed to examine the relationship between code-switching type and age group. The expected and observed frequencies are shown in Table 5. Age influences code-switching patterns, as indicated by significant differences between expected and observed counts.

TABLE 5
OBSERVED VS. EXPECTED FREQUENCIES

Type	16–20 Obs	16–20 Exp	21–30 Obs	21–30 Exp	31–40 Obs	31–40 Exp
Intra-sentential	520	420	500	504	240	336
Tag Switching	250	230	280	276	160	184
Lexical Borrowing	120	180	200	216	220	144
Textisms	150	110	130	132	50	88
Inter-sentential	60	60	40	48	80	72

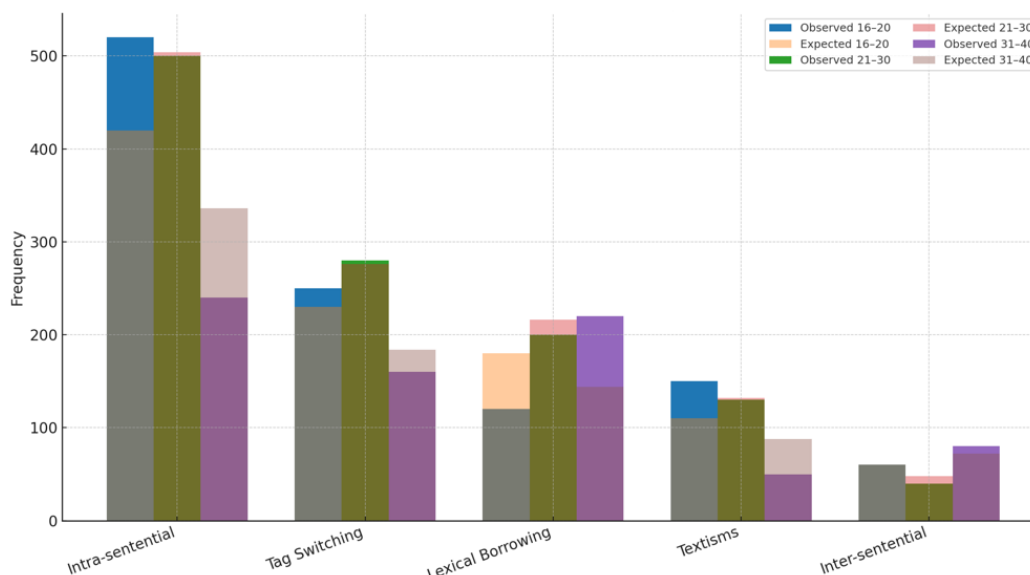


Figure 5. Observed vs. Expected Frequencies of Code-Switching Types by Age Group

In Figure 5 the expected and observed frequencies for each age group are contrasted in the stacked bar chart. The biggest differences are seen in older participants' lexical borrowing and younger participants' intra-sentential switching. The statistical conclusion that code-switching tendencies are age-sensitive is supported by this image.

IV. DISCUSSION

The findings of this research offer important information regarding the traits and functions of code-switching between Arabic and English in SMS interactions. Table 2 and Figure 2 illustrate the frequent occurrence of intra-sentential switching, suggesting a strong degree of bilingual integration among the participants. Previous studies on Spanish–English code-switching indicate that intra-sentential switching demonstrates speakers' proficiency in both languages and their capacity to navigate complex syntactic structures seamlessly. This phenomenon appears to be intensified in the Arabic–English setting due to the widespread use of the English language in social, technological, and educational domains.

Age group analysis identified distinct generational variations in switching behavior (Table 3, Figure 3). Intra-sentential switching and textisms are prevalent in the 16–20 age range, suggesting that younger bilinguals are more familiar with globalized digital settings and more inclined to experiment with hybrid language forms. This reinforces the claim that younger speakers often spearhead online multilingual activities. Conversely, the age range of 31 to 40 demonstrated a distinct inclination towards lexical borrowing. This tendency can be attributed to professional environments where English terminology, especially related to science, technology, and business, is directly included in Arabic discussions without any structural adaptation.

The functional analysis (Table 4, Figure 4) indicates that the need to fill lexical gaps and politeness strategies primarily drive switching. These characteristics indicate that code-switching in SMS is often employed for practical purposes, such as maintaining politeness, achieving clarity, or compensating for vocabulary gaps in one language, rather than merely being a stylistic choice. SMS communication serves as a platform for creating identity, fostering group membership, and promoting cultural hybridity, underscoring the significance of informality and markers of online culture.

Age is an important factor that influences code-switching behaviors, as indicated by the chi-square analysis (Table 5, Figure 5). Younger users frequently switch between sentences, whereas older users utilize vocabulary more than expected, underscoring the discrepancy between actual and anticipated frequencies. This indicates a significant correlation among exposure, education, communicative norms, and generational linguistic practices. This observation supports the socio-pragmatic theory of code-switching as a behavior influenced by social factors rather than solely a linguistic occurrence.

Taking everything into account, the results indicate that code-switching between Arabic and English in SMS communication serves multiple purposes, is contextually dependent, and varies across generations. A complicated bilingual environment, where the selection of language both mirrors and influences evolving digital communication methods, is signified by the interplay of structural types, functional motivations, and demographic variables. These findings not only mirror trends seen in other bilingual contexts but also highlight the unique sociolinguistic features of the Arabic–English combination, particularly within mobile messaging environments.

V. CONCLUSION

This research has analyzed structural patterns, functional motivations, and demographic influences of Arabic–English code-switching in SMS, providing a thorough overview of bilingual digital communication in a Middle Eastern context. Intra-sentential switching emerged as the most common form, demonstrating the strong proficiency of high bilinguals and their seamless ability to merge Arabic and English within individual utterances. Functional analysis revealed that code-switching fulfilled a wide range of communicative functions, including politeness techniques and lexical fillers as main catalysts, as well as stylistic expression and informal indicators consistent with internet culture. Statistical differences between generations were significant, with younger participants (16–20) showing a preference for intra-sentential code-switching and textisms, whereas older participants (31–40) demonstrated more lexical borrowing. The results confirm the influence of generations, educational backgrounds, and social environments on code-switching behavior.

The findings conform to patterns recognized in earlier instances of bilingual communication, like Spanish–English dialogue, and also highlight distinct characteristics of the Arabic–English combination—particularly, the integration of Arabic script, English vocabulary, and digital abbreviations in mixed messages. This fusion of language shows wider trends of globalization, technology integration, and cultural mixtures in modern communication. Through the integration of quantitative frequency analysis, qualitative functional categorization, and the graphical representation of results with figures and tables, the research adds value both theoretically and methodologically to bilingualism and computer-mediated communication. Subsequent investigations could improve this research by employing larger data sets, exploring gender-specific trends, or sampling additional online platforms like WhatsApp and social media to see if the observed patterns are consistent across different communication methods. Finally, Arabic–English code-switching in SMS is shown to be beyond merely a linguistic occurrence; it is an active, contextual activity reflecting identity, practicality, and adaptability in an increasingly globalized world.

ACKNOWLEDGEMENTS

The authors wish to thank Zarqa University. This work was supported by Zarqa University.

REFERENCES

- [1] Alhusna, A., & Febriana, N. (2025). Constructing social identity: A linguistic analysis of generational language use on social media. *EDU Research*, 6(2). <https://iicls.org/index.php/jer/article/view/768>
- [2] Al-Khatib, M. A., & Sabbah, E. H. (2008). Language Choice in Mobile Text Messages among Jordanian University Students. *Finnish Journal of Linguistics*, 21, 37-65.
- [3] Al-Sarayreh, A. (2023). The sociolinguistics of digital communication in Jordan: Language choice and identity on WhatsApp. *Jordan Journal of Modern Languages and Literatures*. <http://journals.yu.edu.jo/jjml/>
- [4] Androutsopoulos, J. (2007). Language Choice and Code-Switching in German-Based Diasporic Web Forums. *The multilingual Internet: Language, culture, and communication online*, 10(4), 524–551.
- [5] Androutsopoulos, J. (2015). Networked Multilingualism: Some Language Practices on Facebook and Their Implications. *International Journal of Bilingualism*, 19(2), 185-205.
- [6] Androutsopoulos, J. (2025). *The Routledge Handbook of Language and Digital Communication* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315754352>
- [7] Barasa, S. (2016). Spoken Code-Switching in Written Form. Manifestation of Code-Switching in Computer Mediated Communication. *Journal of Language Contact*, 9(1), 49-70.
- [8] Common Sense Media. (2023). *Constant companion: A week in the life of a young person's smartphone use*. <https://www.common SenseMedia.org/research/>
- [9] CTIA. (2025). *2025 Annual survey highlights: Americans use record wireless data*. <https://www.ctia.org/news/2025-annual-survey-highlights>
- [10] Cunningham, K. J. (2024). Translingualism and digital communication: New perspectives on multilingual identity. *International Journal of Bilingualism*. <https://doi.org/10.1177/13670069241234567>
- [11] Dascalu, E. A. (1999). The Form and Function of Code-Switching in Bilingual Writing: A Comparative Analysis of English Email Messages Written by Belarusian/English, Korean/English, and Romanian/English Bilinguals. *The University of Tulsa Graduate Review*. <http://www.utulsa.edu/tugr/tugr99/codeswit.html>
- [12] Dorleijn, M., & Nortier, J. (2008). *Code-Switching and the Internet*. Cambridge: Cambridge University Press.
- [13] Elhija, D. A. (2023). Code switching in digital communication: Arabic and Hebrew. *Open Journal of Modern Linguistics*, 13(3), 357–376. <https://doi.org/10.4236/ojml.2023.133021>
- [14] Fernández-Mallat, V. (2020). “Cuando Me da la Gana. Me AF”: Washingtonian Bilingual Speakers of Spanish on Facebook. In *Spanish across Domains in the United States* (pp. 289-311). Brill.
- [15] Georgakopoulou, A. (1997). Self-Presentation and Interactional Alliances in E-Mail Discourse: The Style and Code-Switches of Greek Messages. *International Journal of Applied Linguistics*, 7(2), 141-164.
- [16] Goldbarg, R. N. (2009). Spanish-English Codeswitching in Email Communication. *Language@ Internet*, 6(3), 1–21.
- [17] Halim, N. S., & Maros, M. (2014). The Functions of Code-Switching in Facebook Interactions. *Procedia-Social and Behavioral Sciences*, 118, 126-133.
- [18] Hinrichs, L. (2006). *Codeswitching on the Web*. The Netherlands & Philadelphia, PA: Benjamins.
- [19] Hinrichs, L. (2016). Modular Repertoires in English-Using Social Networks: A Study of Language Choice in the Networks of Adult Facebook Users. *English in computer-mediated communication: Variation, representation, and change*, 93, 17.
- [20] Iqbal, S. (2025). Digital dialects: Code-switching patterns in social media communication. *Global Sociological Review*, X(II), 83–98. [https://doi.org/10.31703/gsr.2025\(X-II\).07](https://doi.org/10.31703/gsr.2025(X-II).07)
- [21] Jurgens, D., Dimitrov, S., & Ruths, D. (2014). Twitter Users #Codeswitch Hashtags! #Moltoimportante #wow. In *Proceedings of the First Workshop on Computational Approaches to Code Switching* (pp. 51-61). Association for Computational Linguistics.
- [22] Kytölä, S. (2013). *Multilingual language use and metapragmatic reflexivity in Finnish internet football forums: A study in the sociolinguistics of globalization*. University of Jyväskylä.
- [23] Lavender, J. (2017). Comparing the Pragmatic Function of Code Switching in Oral Conversation and in Twitter in Bilingual Speech from Valencia, Spain. *Catalan Review*, 31, 15-39.
- [24] Lee, C. (2016). *Multilingualism online*. Routledge.
- [25] Leppänen, S. (2012). Linguistic and Generic Hybridity in Web Writing: The Case of Fan Fiction. In *Language mixing and code-switching in writing* (pp. 233-254). Routledge.
- [26] Lee, C. (2023). *Multilingualism online: Language, identity and digital communication*. Routledge. <https://doi.org/10.4324/9781003328452>
- [27] Leppänen, S., & Peuronen, S. (2012). Multilingualism on the Internet. In *The Routledge handbook of multilingualism* (pp. 398-416). Routledge.
- [28] Liénard, F., & Penloup, M.-C. (2011). Language Contacts and Code-Switching in Electronics Writings: the Case of the Blog. In *Code-switching, Languages in contact and Electronic Writings* (pp. 73-86). Peter Lang.
- [29] Lin, A. (2005). Gendered, Bilingual Communication Practices: Mobile Text-Messaging among Hong Kong College Students. *Fibreculture Journal*.
- [30] Magaña, D. (2013). Code-Switching In Social Network Messages: A Case Study of a Bilingual Chicana. *International Journal of the Linguistic Association of the Southwest (IJLASSO)*, 32(1), 43–65.
- [31] Montes-Alcalá, C. (2005). ¡ Mándame un Email!: Cambio de Códigos Español-Inglés Online. *Contactos y contextos lingüísticos: el español en los Estados Unidos y en contacto con otras lenguas*-(Lingüística iberoamericana; 27), 173-185.
- [32] Montes-Alcalá, C. (2007). Blogging in Two Languages: Code-Switching in Bilingual Blogs. *Selected proceedings of the third workshop on Spanish sociolinguistics*.
- [33] Montes-Alcalá, C. (2016). iSwitch: Spanish-English Mixing in Computer-Mediated Communication. *Journal of Language Contact*, 9(1), 23-48.
- [34] Montes-Alcalá, C. (2024). Bilingual texting in the age of emoji: Multilingualism in SMS. *Languages*, 9(4), 136. <https://doi.org/10.3390/languages9040136>

- [35] Novianti, W. (2013). The Use of Code Switching in Twitter (A Case Study in English Education Department). *Passage*, 1(2), 1-10.
- [36] Paolillo, J. (2011). "Conversational" Codeswitching on Usenet and Internet Relay Chat. *Language@ Internet*, 8.
- [37] Paolillo, J. C. (1996). Language Choice on soc.culture.punjab. *Electronic Journal of Communication/La revue électronique de communication*, 6(3).
- [38] Pérez-Sabater, C. (2022). Mixing Catalan, English and Spanish on WhatsApp: A Case Study on Language Choice and Code-Switching. *Spanish in Context*, 19(2), 289-313.
- [39] Sabty, C. (2024). Computational and sociolinguistic approaches to Arabic-English code-switching. *Proceedings of the 2024 Joint International Conference on Computational Linguistics*. <https://aclanthology.org/2024.lrec-main.102>
- [40] Ting, S.-H., & Yeo, D. K.-L. (2019). Code-Switching Functions in Facebook Wallposts. *Human Behavior, Development & Society*, 20(3), 7-18.
- [41] Tsiplakou, S. (2009). Doing (Bi) Lingualism: Language Alternation as Performative Construction of Online Identities. *Pragmatics Quarterly Publication of the International Pragmatics Association (IPrA)*, 19(3), 361-391.
- [42] Vallejo, L. L. (2011). El Cambio de Código Español-Inglés como Creatividad Lingüística y Presentación de la Imagen en Tweets Escritos por Tijuaneños. *Memorias en extenso de las Primeras Jornadas de Lenguas en Contacto*, 64-73.
- [43] Verheijen, L. (2019). *Is textese a threat to traditional literacy? Dutch youths' language use in written computer-mediated communication and relations with their school writing*. Utrecht: LOT.
- [44] Verheijen, L., de Weger, L., & van Hout, R. (2018). Code-Mixing with English in Dutch Youths' Online Language: OMG SUPERNICE LOL. In Reinhild Vandekerckhove, Darja Fišer & Lisa Hilde (Eds.), *Proceedings of the Sixth Conference on Computer-Mediated Communication (CMC) and Social Media Corpora*. Antwerpen: University of Antwerp.
- [45] Verheijen, L., & Van Hout, R. (2022). Manifold Code-Mixing in Computer-Mediated Communication: The Use of English in Dutch Youths' Informal Online Writing. *Ampersand*, 9, 100091.
- [46] Warschauer, M., Said, G. R. E., & Zohry, A. G. (2002). Language Choice Online: Globalization and Identity in Egypt. *Journal of Computer-Mediated Communication*, 7(4), JCMC744.
- [47] Wentker, M. (2018). Code-Switching and Identity Construction in WhatsApp. *The Discursive Construction of Identities On- and Offline*, 109-131.

Wasan Khalid Ahmed was born in Baghdad, Iraq, in 1974. She holds a Ph.D. in English Language Studies/Applied Linguistics from the Faculty of Social Sciences & Humanities, Universiti Kabangsaan Malaysia (UKM). She is currently a lecturer at the Department of English Language and Literature at Fallujah University, Iraq. Her research interests span applied linguistics, ESP, and pragmatics. She explores discourse studies, politeness, and cultural influences. Additionally, she investigates media's role and gender dynamics in communication. Her work highlights the complex interplay of language and social identity. Email: Wasan.khalid@uofallujah.edu.iq; <https://orcid.org/0000-0002-7137-390X>

Ibtihal Hassan Mussa was born in Karbala, Iraq. She is currently a Ph.D. candidate at the Faculty of Educational Sciences and Technology in University of Technology Malaysia, Johor Bahru, Malaysia. She is working as school teacher in Iraq. Her research interests span technology in education and linguistics. Email: ibtihal.hassan@ymail.com; <https://orcid.org/0009-0002-6509-0812b>



Laith Salman Hadla was born in Baghdad, Iraq in 1970. He was awarded his PhD Degree in Translation and Linguistics from Al Muatansiriya University, Baghdad in 2006. He is an Associate Professor of Translation and Linguistics at Zarqa University / Jordan. His research interest includes different fields of Interpretation, Translation, and Linguistics. He is a member in Jordanian Translators and Applied Linguists Association and Jordanian Translators Association. He can be contacted at: lhada@zu.edu.jo



Marwan Harb Alqaryouti was born in Zarqa, Jordan in 1969. He was awarded his PhD Degree in English Language Studies/American Literature from the Faculty of Languages and Communication at Universiti Sultan Zainal Abidin in Terengganu, Malaysia. He is currently an assistant professor at Zarqa University / Jordan. His research interest includes different fields of English Literature. Dr. Alqaryouti is a member of the Jordanian Translators and Applied Linguists Association and Jordanian Translators. He can be contacted at: mqaryouti@zu.edu.jo



Ala Eddin Sadeq was born in Mafraq, Jordan in 1971. He was awarded his PhD Degree in English Literature from the University of Rajasthan, India in 2000. He is a Professor of English Literature at Zarqa University / Jordan. His research interest includes different fields of English Literature. Prof. Ala Eddin Sadeq is the Vice President of Zarqa University for Academic Affairs. Professor Ala Eddin is the Secretary General of the English Language International Conference - ELIC. He is a member in Jordanian Translators and Applied Linguists Association and Jordanian Translators Association. He can be contacted at: alaeddin71@yahoo.com

Documenting Linguistic Diversity: An Analysis of Compound Word Formation in the Goris Armenian Dialect

Lusine R. Nersisyan

Philology and Historical Jurisprudence, Goris State University, Goris, Armenia;
Syunik Center for Armenological Studies, Goris, Armenia

Mher S. Kumunts

Dialectology Department, Institute of Language named after Hrachya Acharyan of NAS RA, Yerevan, Armenia;
Philology and Historical Jurisprudence, Goris State University, Goris, Armenia

Ani A. Ghazaryan

Faculty of Armenian Language and Literature, Goris State University, Goris, Armenia

Manuk H. Falakyan

Western Armenian Language Department, Institute of Language named after Hrachya Acharyan of NAS RA, Yerevan,
Armenia;
Yerevan State University, Yerevan, Armenia

Armine A. Mkrtumyan

Department of Lexicography, Institute of Language named after H. Acharyan of NAS RA, Yerevan, Armenia;
Department of Armenian and Foreign Languages, Armenian National Agrarian University, Yerevan, Armenia

Leonid G. Telyan

Department of the Modern Armenian Language, Institute of Language named after Hrachya Acharyan of NAS RA,
Yerevan, Armenia

Anahit S. Galstyan

Department of Lexicography, Institute of Language named after H. Acharyan of NAS RA, Yerevan, Armenia

Susanna K. Tioyan

Western Armenian Language Department, Institute of Language named after Hrachya Acharyan of NAS RA, Yerevan,
Armenia

Harout K. Marachlian

Western Armenian Language Department, Institute of Language named after Hrachya Acharyan of NAS RA, Yerevan,
Armenia;
American University of Armenia, Yerevan, Armenia

Tereza S. Shahverdyan

Armenian Language and Literature, Vanadzor State University, Vanadzor, Armenia

Anna H. Abajyan

History of Armenian Language and General Linguistics, Faculty of Armenian Philology, Yerevan State University,
Yerevan, Armenia

Manana G. Dalalyan

English Philology, Faculty of European Languages and Communication, Yerevan State University, Yerevan, Armenia

Vanuhi M. Baghrumyan

Lexicography Department, Language Institute named Hrachya Acharyan of the National Academy of Sciences of the
Republic of Armenia, Yerevan, Armenia;
Base High School, Khachatur Abovyan Armenian State Pedagogical University, Yerevan, Armenia

Karine L. Arakelyan

Modern Armenian Language Department, Institute of Language named after Hrachya Acharyan of NAS RA, Yerevan, Armenia

Abstract—This study investigates the structural, semantic, and phonetic peculiarities of genuine compound words in the Goris dialect, a vital component of the rich Armenian dialectal system. In an era of globalization where dialects face the risk of extinction due to the strong influence of literary languages, their documentation and scientific description acquire significant cultural and historical value. Employing descriptive, comparative-contrastive, and structural analysis methods, the research meticulously examines Goris dialectal phenomena by contrasting them with Literary Eastern Armenian and the closely related Artsakh dialect, and drawing upon established theories of Armenian word formation. Genuine compound words in the Goris dialect, being synthetic-type formations, on the one hand, follow the general word-formational regularities of Modern Armenian. The findings reveal that Goris compound words, while adhering to general Armenian word-formational principles (e.g., noun+noun, adjective+noun, noun/adjective+verbal root models, with and without connecting vowels), exhibit distinctive phonetic characteristics. A prominent feature is the consistent transformation of the connecting vowel 'a' to 'ə' (ə) in unstressed positions, imparting a unique phonetic fingerprint to the dialect. The study also identifies exceptional formations, including compounds derived from idiomatic structures. A comparative analysis with the Artsakh dialect highlights shared lexical and structural foundations, underscoring their genetic proximity. These findings underscore the dynamic and independent nature of dialects, contributing to a deeper understanding of Armenian linguistic history and the broader field of dialectology and language preservation.

Index Terms—Goris dialect, compound words, Armenian linguistics, dialectology, language preservation

I. INTRODUCTION

The Goris dialect, as an integral part of the rich Armenian dialectal system, is a linguistic unit that requires thorough examination. In the context of globalization, where dialects are exposed to the strong influence of the literary language and face the risk of extinction, their documentation and scientific description acquire not only linguistic but also cultural and historical value. As Vaux (2024) powerfully demonstrates in his recent comprehensive overview, many non-standard Armenian dialects are in imminent danger of disappearing, a situation that underscores the critical urgency for their thorough documentation and nuanced linguistic analysis. This call for intensified effort in documenting Armenian dialects has also been highlighted by other international scholarship (Weitenberg, 2002). Indeed, such critical urgency for thorough documentation and nuanced linguistic analysis is widely recognized as vital in the broader field of language endangerment and preservation (Austin & Sallabank, 2011).

The Goris dialect is a prominent Eastern Armenian dialect, geographically situated in the Syunik province of Southern Armenia, which borders the Islamic Republic of Iran and the Republic of Azerbaijan. It also historically bordered Artsakh. It belongs to the 'um' branch of Armenian dialects, specifically within the Syunik-Artsakh dialect group. A defining characteristic of the Goris dialect is its distinctive phonetic system, most notably the consistent transformation of the connecting vowel 'a' to 'ə' (schwa) in unstressed positions, which lends a unique phonetic color to its pronunciation (Vaux, 2008; Vaux & Suhairi, 2024). Morphologically, it adheres to the 'um' type present tense formation, a hallmark of Eastern Armenian varieties. Furthermore, the dialect is known for its preservation of archaic vocabulary and a rich array of compound words, including many derived from idiomatic structures, underscoring its dynamic and historically significant linguistic system (Kumunts, 2022; Margaryan, 1976).

This study aims to highlight the uniqueness of the Goris dialect and to preserve its lexical richness for future generations. A dialect is a crucial testament to the existence of a people and its history, always requiring scrutiny and, even more so, preservation.

The main objective of this study is to comprehensively examine the structural, semantic, and phonetic peculiarities of genuine compound words in the Goris dialect, contrasting them with Literary Eastern Armenian and the Artsakh dialect. This approach aligns with contemporary calls in Armenian dialectology to expand research beyond purely historical and prescriptive issues, incorporating more synchronic grammatical and sociolinguistic analysis (Vaux, 2024).

To achieve this objective, the following tasks have been set:

To present the theoretical foundations of genuine compound words, based on the works of Armenian linguists.

To classify the genuine compound words of the Goris dialect according to their formation (with connecting vowel, without connecting vowel).

To analyze the role of the connecting vowel (particularly 'a' and its phonetically altered variant 'ə') in dialectal word formation.

To conduct a comparative analysis of genuine compound words in the Goris and Artsakh dialects, revealing their commonalities and differences.

To demonstrate the phonetic and structural differences in the formation of compound words between the dialect and the literary language.

The comprehensive examination of the Goris dialect was undertaken by Aleksandr Margaryan in his monograph "The Goris Dialect," where he also addressed genuine compound words (Margaryan, 1976). The place and role of the dialect within the Armenian dialectal system have been studied by Kumunts (2022), and further perspectives on Armenian dialectology are provided by international scholars like Weitenberg (2002), Greppin and Khachaturian (1986), and Katvalyan (2014).

However, to examine the word-formational phenomena of the Goris dialect in a broader Armenian linguistic context, it is crucial to rely on the general theory of compounding and word formation. For Modern Armenian, the two main manifestations of word formation patterns—derivation and compounding—have been fundamentally defined by Jahukyan (1989). Expanding on this, Jahukyan (1989), in examining the general types of word formation patterns, distinguishes coordinative and subordinate compounds, characterizing them by such features of substantiality as composition, number, dependency, and class. Of these criteria, the first two (composition and number) relate to compounding as an integral structure, while the latter two (dependency and class) pertain to the interrelationship of its components. Indeed, the field of word-formation, encompassing derivation, compounding, and conversion, has seen a significant revival in linguistic study since the mid-20th century, becoming a central area of research (Štekauer & Lieber, 2005). In this regard, a series of works by Armenian linguists have served as a foundation. Specifically, Eloyan's (1970) article dedicated to compound words with and without connecting vowels provides an opportunity to juxtapose the regularities of the literary language with dialectal realities. Mkrtchyan's (1972) proposed classification of compound words allows for a critical analysis of traditional approaches. From the perspective of the theory of word-formational patterns, Margaryan's (1990) observations are valuable, and from the historical origin of compound words, Khachaturyan's (2020) insights are significant. The synthesis of the theoretical provisions of these works allows for the examination of genuine compound words in the Goris dialect not in isolation, but as an original manifestation of the overall word-formational system of the Armenian language.

The research was conducted using several linguistic methods. The primary method employed was the descriptive method, through which the genuine compound words of the Goris dialect were presented and analyzed. The comparative-contrastive method was also applied to juxtapose dialectal phenomena with data from the literary language and the Artsakh dialect, aligning with established practices in language variation studies (Chambers et al., 2021). When analyzing the internal structure of genuine compound words, the method of structural analysis was utilized, reflecting current best practices in dialectal research (Vaux, 2024). The materials for the study included word lists available in Margaryan's (1976) work, as well as words and expressions recorded from dialect speakers by the authors during fieldwork.

II. TYPES OF COMPOUND WORDS AND THE PLACE OF GENUINE COMPOUND WORDS IN THE WORD-FORMATIONAL SYSTEM

In enriching the vocabulary of modern Armenian, compounding, along with affixation (derivation), is one of the most effective methods, as noted by Aghayan (1984). In Armenian linguistics, various approaches exist regarding the classification of compound words. Traditionally, a distinction is made between genuine (proper) and non-genuine compound words. However, as R. Mkrtchyan notes in his study, this division is not always consistent, as the characteristics typical of both types often overlap (Mkrtchyan, 1977). Therefore, he proposes dividing compound words into two main types: *synthetic* (compositional) and *analytic*. Synthetic compounds are those whose components "appear as a unified whole and possess a single lexical stress" (Mkrtchyan, 1972). The subject of our study, genuine compound words, belongs to the synthetic type according to this classification.

Essentially, genuine compound words are words formed by the combination of two or more roots, which express a new, complete meaning. This aligns with broader morphological theories (Booij, 2018). As observed by L. Hovsepyan, the connection between their components is primarily subordinate, meaning one component modifies the other (Hovsepyan, 2016).

According to the means connecting the components, genuine compound words are of two types: with a connecting vowel and without a connecting vowel.

A. Compound Words With a Connecting Vowel

The main connecting vowel in Armenian is *a*. S. Eloyan, in his article, thoroughly examines the role of the connecting vowel, noting that it, "having no semantic participation, plays a purely morphological role" and cannot be considered a full-fledged morpheme (Eloyan, 1970). In the literary language, as the author details, a stable regularity operates: when the first component ends with the vowel *i*, it becomes *e* when combined with the connecting vowel *a* (*i+a > e*), for example, *այգի+գործ* → *այգեգործ* [*aygi+gorts* → *aygegorts*] 'gardener', *բարի+կամ* → *բարեկամ* [*bari+kam* → *barekam*] 'friend'. Knowledge of this regularity is important for evaluating the peculiarities of the Goris dialect.

B. Compound Words Without a Connecting Vowel

The formation of compound words without a connecting vowel, as Khachaturyan notes, citing Jahukyan, is an older, Indo-European heritage (Khachaturyan, 2020). Compound words without a connecting vowel are primarily formed when the second component begins with a vowel (e.g., *բազմերանգ* [*bazmerang*] 'multi-colored', *հաւրօգութ* [*hanrōgut*]

'publicly beneficial'), because, as Eloyan substantiates, the connecting vowel usually drops out before a vowel (Eloyan, 1970).

III. STRUCTURAL-SEMANTIC EXAMINATION OF GENUINE COMPOUND WORDS IN THE GORIS DIALECT

The Goris dialect, while conforming to the general word-formational regularities of Armenian, at the same time exhibits a number of unique features in the formation of genuine compound words. In the dialect, as in the literary language, genuine compound words can be with a connecting vowel and without a connecting vowel (Margaryan, 1976).

A. Genuine Compound Words With a Connecting Vowel and Their Phonetic Peculiarities

In the Goris dialect, as in Literary Armenian, genuine compound words with a connecting vowel constitute a predominant number (Margaryan, 1990). However, in contrast to the literary language, where the connecting vowel *a* is either stable or transforms into *e*, in the Goris dialect, it demonstrates exceptional phonetic flexibility, which is one of the most expressive features of the dialect.

According to Margaryan's classification and our observations, the following manifestations of the behavior of the connecting vowel *a* can be distinguished (Margaryan, 1976):

(a). The Connecting Vowel 'a' Remains Unchanged

It remains unchanged when it is stressed, mainly in the case of a monosyllabic final component: քրշ+ա'+թօխս [k'ěsh+á+t'ókhp] 'fog', ծրք+ա'+փօղ [tsěrk'+á+p'ógh] 'hand-pipe (a specific type of pipe)', քրթ+ա'+ծակ [k'ět'+á+tsak] 'nostril, figuratively: intelligence, wit'.

(b). The Connecting Vowel 'a' Transforms Into 'ä'

This occurs due to the law of vowel harmony (mainly under the influence of *i* and *e* vowels), again appearing in compound words with a monosyllabic final component: իրի+ա'+տնղ [iris+ä+tēgh] 'respect, honor', ճիր+ա'+ծէտ [chir+ä+tsēt] 'a type of bird found on riverbanks', վիլ+ա'+տնի [vil+ä+tü] 'spinning wheel tool'.

(c). The Connecting Vowel 'a' Transforms Into 'ě' ('ə')

This is the most characteristic and widespread phonetic regularity of the dialect, occurring when the connecting vowel appears in an unstressed position before the stressed syllable. This phenomenon is observed in words with both polysyllabic and monosyllabic final components and is considered the phonetic hallmark of the dialect: ծնն+ը+խառնը [tsün+ě+khár'ně] 'snow-mixed', վնդ+ը+կաղի [vënd+ě+kaghī] 'foot-stuck/footed', սրբ+ը+կուտուր [sěrt+ě+kutur] 'heartbroken', քաղր+ը+կին [ěkhpěr+ě+kin] 'brother's wife', ըշկ+ը+տիքա՞ծ [ěshk+ě+tiräts] 'having an eye on / observed'.

(d). Other Phonetic Transformations of the Connecting Vowel 'a'

In some cases, again under the influence of the law of vowel harmony, *a* can also transform into *i* or *u/ü* vowels, such as ճիր+ի+թաթախ [chir+i+t'at'akh] (or ճիրըթաթախ [chir'ět'at'akh]) 'water-soaked', լից+ի+կնւմէլ / լնց+նի+կնւմէլ [lits'+i+kümēl / lüts'+ü+kümēl] 'to cry'.

(e). In Compound Words Borrowed From the Literary Language

It is interesting that in compound words adopted by the dialect from the literary language, the connecting vowel, as a rule, does not undergo dialectal phonetic transformation and remains *a*, although it may be reduced in pronunciation: ըղ+ա+ճնր [ěgh+a+chür] 'salt water', լղ+ա+ճաք [lēgh+a+chak] 'fear', վնդ+ա+կօխ [vënd+a+kōkh] 'trampled by foot'.

B. Genuine Compound Words Without a Connecting Vowel

In contrast to compound words with a connecting vowel, genuine compound words without a connecting vowel are significantly fewer in number in the Goris dialect, as in the literary language. They are formed mainly in two cases, following the general word-formational regularities of Armenian (Margaryan, 1976).

(a). When the Second Component Begins With a Vowel

This is the most common model for the formation of compound words without a connecting vowel, which, as mentioned in the theoretical section, is due to the omission of the connecting vowel before a vowel.

Examples: սըվ+ի'րէս [sěv+irēs] 'black-faced', ցրի+ւ'իլ [ts'ěkh+évil] 'broom', մըթ+ա'ւտղ [mět'+ásēgh] 'pin', յըղ+ա'ման [yěgh+áman] 'oil-lamp/oil-jar'.

(b). When the Second Component Begins With a Consonant

This group is smaller and more interesting because here the absence of a connecting vowel is not determined by a phonetic law, but by word-formational tradition. Often, words in this group are formed by combining a noun and a verbal root or a word expressing an action.

Examples: հաղ+թօխ [hats'+t'ókhh] 'baker', թէր+խաշ [t'ēr+khash] 'partially cooked, half-cooked', ճրքք+կալ [chěřk'+kal] 'lamp stand/candelabra', քթ+կալ [k'ět'+kal] 'woman's facial veil'.

(c). *Other Peculiarities of Compound Words Without a Connecting Vowel*

Despite their small number, genuine compound words without a connecting vowel in the Goris dialect exhibit some interesting manifestations:

1. *Historical Omission of the Connecting Vowel*: Some words that, in the past, were likely formed with a connecting vowel have lost it in the dialect and are now used as compound words without a connecting vowel. This demonstrates the dynamic development of the language. For example, the Literary Armenian compound word նախաթթու [chakhat't'u] 'beetroot pickle', which includes a connecting vowel, has become the compound word ճրխթթո [chëkht'ët'ō] without a connecting vowel in the Goris dialect. Similarly, թթու+խմոր [t't'u+khmor] 'sourdough' > տրխտթոր/թրթխտթոր [tëkhtëmör/t'ët'khëmör] 'sourdough'.
2. *Connection with Idiomatic Structures*: Some compound words without a connecting vowel, in their structure, resemble idioms or stable collocations that have, over time, fused and become a single word. For instance, the word չափարախ/չմփարախ [ch'ap'arakh/ch'ëmp'arakh] 'gallop (four-hoofed)' (meaning 'four-hoofed gallop') may have originated from the combination չորս+nup [ch'ors+otk] 'four+legs'.
3. *Exceptional Formations*: The compound word դիբիրգադեմ [dibirgädēm] (դեպի+երեկոյ+ա+դեմ [depi+yerekoy+a+dēm] 'towards+evening+a+face/front' meaning 'towards evening') is exceptional not only due to the presence of a connecting vowel (which deviates from the general rule of this section) but also by its formation, which includes a preposition and multiple components. However, considering it as a fusion of the stable collocation "դեպի երեկո" [depi yereko] 'towards evening', it represents a unique type of compound derived from an idiomatic structure.

Another remarkable example of a unique formation, deeply rooted in the historical development of Armenian, is the word սարսընգի || արսընգի [sarasēngi || arasēngi], used in the dialectal area to mean "swing" or "a long wooden swing." This term, primarily used in children's games ("Սարս-սընգի, արս-սընգի, քարին իրա՞ն տախտակը տնգի" [sara-sēngi, sara-sēngi, k'arin irān taxtakē tēngi] "Swing, swing, put the board on the stone"), is attested in the Syunik-Artsakh dialect group, including Urmia (Sargsyan, 2008). Its etymology traces back to the Proto-Indo-European root **sueng(k)*- meaning 'to bend, curve, swing.' This root has phonetic and semantic parallels in Germanic dialects (e.g., German *schwenken* 'to swing', Old High German *swenken* 'swing'), as well as in Old Irish (*seng* 'thin, slender') and Old Saxon (*swingan* 'to swing oneself') (Torosyan, 2018; Pokorny, 1959). The variant արսընգի [arasēngi] is believed to have undergone a phonetic shift (**sueng(k)* > *ar-* *sueng(e)*) due to compounding, demonstrating a different phenomenon from the regular Proto-Indo-European *s* > *h* || 0 (*s* > *h* or 0) change in Armenian (cf. առ-(ա)-սպել [ar-(a)-spe] < Gothic *spill* 'story') (Jahukyan, 1987). The form սարսընգի [sarasēngi] is thought to have developed within the dialect group through analogy with the initial u- [s-] of the second component. This word is not found in early literary sources or linguistic studies, highlighting its ancient, purely dialectal origin and significance for understanding the deeper historical layers of the Goris dialect.

C. *Word-Formational and Semantic Analysis of Genuine Compound Words*

The examination of genuine compound words in the Goris dialect would not be complete without an analysis of the parts of speech and semantic connections of their components. Among the most active patterns in neologisms are the noun+noun, adjective+noun, and verb+noun models, which is fully confirmed by the material of the Goris dialect as well. A statistical analysis of our studied material shows that compound words formed with a connecting vowel predominate in the dialect, and among these, the phonetically altered connecting vowel 'ë' (ə) is, in turn, dominant.

A notable example illustrating the semantic and etymological complexity in the Goris dialect is the compound word գիլիմասրի [gilimasri] 'potato' (a dialectal term for topinambour/ground apple). As Kumunts extensively details, this term, alongside variants like գլլըմասրի [gëlëmasri], is considered a dialectal form derived from the Old Armenian գերմաստ [germast]. Kumunts traces its etymology to the combination of Proto-Indo-European roots **ger-* ('flexible, curved') and **mast* ('thick, solid, firm'), suggesting a semantic evolution that reflects both the plant's physical characteristics and potential historical linguistic shifts within the dialect (Kumunts, 2019). This specific analysis provides valuable insight into the dialect's vocabulary formation and its deep historical layers.

Below, we present examples of genuine compound words characteristic of the Goris dialect, classified according to the most common word-formational models (see Table 1).

layers and linguistic dynamism within this regional variety of Armenian. Crucially, in the challenging context following the Armenian-Azerbaijani conflict (2020-2023), which has severely impacted the native environment of the Artsakh dialect, the Goris dialect's preservation of such unique vocabulary serves as a vital testament to linguistic diversity.

(a). *Etymological and Semantic Origins*

The term փրքսըր [p'ěrsěvēr] is the dialectal equivalent of Literary Eastern Armenian *berank'sivayr* 'face down, prone,' which in turn originates from the phrase բերանս ի վայր [berans i vayr] 'face downwards' (Acharyan, 1971). It is composed of the root փրք [p'ěrs-] (from բերան [beran] 'mouth/face'), the preposition ի [i], and վայր [vayer] 'down/place.' In Proto-Indo-European (PIE) root patterns, the word բերան [beran] 'mouth' (Jahukyan, 2010; Pokorny, 1959) originates from the root **bher-*, whose primary meanings of '1. throat, gullet, 2. narrow passage' are no longer preserved in the Syunik-Artsakh dialect group. Nevertheless, the root փրք [p'ěrs] does not appear independently with the meaning 'mouth/face' either in isolation or as part of other compounds in this dialect, leading to the hypothesis that the suffix '-an' of բերան [beran] developed later, and the dialectal variant փրք [p'ěrs] is simply another form of the common Armenian word բերան [beran].

(b). *Phonetic Evolution*

The phonetic shifts in փրքսըր [p'ěrsěvēr] in the Goris dialect reflect characteristic patterns of the Syunik-Artsakh dialect group (Margaryan, 1976).

1. The initial 'b' of բերան [beran] 'mouth' regularly undergoes lenition, becoming 'p' in the dialect (e.g., բերան [beran] > պերան [p'ėran], բուկ [buk] > պոկ [p'ok] 'mouthpiece', բահ [bah] > պահ [p'ah] 'spade').
2. Concurrently, /b/ often undergoes aspiration, becoming /p'h/ (written as փ) when preceding 'r' or 'r' (e.g., բուրվառ [burvar] 'censer' > փրքվառ [p'hırvār], բռնել [brnel] 'to hold' > փրքնել [p'hırne] 'to catch', Classical Armenian բրդեմ [brdem] 'to pluck' > փրքթիլ [p'hırt'il]). Variant forms like փիրթիլ [p'hirt'il] (with i) suggest an earlier phonetic stage. Other phonetic variants of p'ěrsěvēr, such as փրքսըր || փրքսըր || փիսիքսըր || փրքսըր [p'ěsvēr || p'ěrsěēr || p'isivēr || p'ěrsěyēr], further corroborate these phonological tendencies.
3. The u 's' in p'ěrsěvēr (from բերանս ի վայր [berans i vayr]) originates from the Classical Armenian accusative plural suffix -s, which imparted a generalized meaning to the word, signifying 'the mouths' (e.g., պիր-նին [pir-nin] 'the mouths').
4. The preposition 'ի' [i] (in բերանս ի վայր [berans i vayr]) is frequently reduced to ě (schwa) in most dialectal variants due to stress loss.
5. Finally, the Classical Armenian word վայր [vayr] 'down' evolved into վեր || յեր [vēr || yēr] in the Syunik-Artsakh dialect group. The initial /v/ of վեր [vēr] further shifts to /y/ when preceding /ē/ (e.g., Classical Armenian վերայ [vēray] 'above' > յիրա [yira], վեր [vēr] > յեր [yēr] 'up').

(c). *Semantic Development and Usage*

The diverse applications of p'ěrsěvēr in the Goris dialect demonstrate its semantic dynamism and richness. These include:

Փրքսըր անել [p'ěrsěvēr anel] 'to turn over'

Փրքսըր ինել [p'ěrsěvēr inel] '1. to fall face down, 2. fig. to become ill, 3. (Artsakh dial. (Sargsyan, 2013)) to become annoyed'

Փրքսըր քիցել [p'ěrsěvēr k'itsel] 'to turn'

Փրքսըր յեշել [p'ěrsěvēr yēshel] 'to look askance'

Փրքսըր քոն ինել [p'ěrsěvēr k'on inel] 'to sleep face down'

Փրքսըր վերել [p'ěrsěvērvel] '1. to resist, 2. to fall, 3. to become ill, 4. to die'

(d). *Sociolinguistic and Cultural Context*

The narratives surrounding p'ěrsěvēr, recounted by Elmira Eylazyan (38 years old) and Sveta Snetsunts (67 years old) from Tegh-Khndzoresk, highlight its cultural and historical significance. One account states that 'beautiful girls born in Tegh village were put փրքսըր [p'ěrsěvēr] so they wouldn't fall into the hands of foreigners.' Another version suggests that 'girls born in Tegh village were put փրքսըր [p'ěrsěvēr] so that more boys would be born.' While these stories are products of folk etymology, they reflect the dialect's vitality and its deep connection to local identity and historical circumstances.

F. *Etymological Analysis of the Reduplicative Compound կուրկուր [Kurkup]: A Comparative View*

Another compelling example of the Goris dialect's lexical depth is the reduplicative compound *kurkup* (կուրկուր), which carries meanings related to being silent, huddled, or closed off. The etymology of this word is subject to debate, with two primary hypotheses offering different explanations for its formation. The following table provides a comparative analysis of these two viewpoints, examining the proposed structure, origin of components, and supporting evidence for each.

TABLE 3
COMPARATIVE ETYMOLOGICAL HYPOTHESES FOR THE DIALECTAL WORD KURKUP'

Criterion	Hypothesis 1: Compound of Two Distinct Roots	Hypothesis 2: Reduplication of a Single Root
Core Principle	The combination of two different Proto-Indo-European (PIE) roots.	The reduplication and subsequent phonetic alteration of a single root.
Word Structure	կուր- [kur-] (separate root) + կուի- [kup'-] (separate root)	կուր- [kur-] (altered form of -կուի [-kup']) + կուր- [kur-].
Origin of First Component (+ կուր- [kur-])	Derived from PIE root *ghūbh-. Slavic cognates: Russian gubá (губа, 'lip'), Serbian gübica ('muzzle'), Bulgarian gъba (гъба, 'mouth'). Other cognates: Greek gamphai (γαμφαί, 'jawbone').	An altered, reduplicated version of the root -կուի [-kup'].
Origin of Second Component (- կուի [-kup'])	Derived from PIE root *ghaph- ('to grab, seize'). Based on the root կաի [kap'] from the verb կաիւնցաւնել [kap'uts'anel] ('to seal, close'), as well as կաի [kap'], կուի [khuip'].	Agrees with Hypothesis 1: derived from PIE root *ghaph- ('to grab, seize') → կուի [khuip'].
Formation Mechanism	Simple compounding of the two proposed roots.	Reduplication with dissimilation: կուի-կուի [kup'-kup'] → Regressive dissimilation of the initial /p/ (փ) to /t/ (ք) → կուրկուի [kurkup']. Parallels in Armenian: թափ-թափ [t'ap'-t'ap'] > թաղթաղ [t'art'ap'] ('drizzle'); սափ-սափ [sap'-sap'] > սարսափ [sarsap'] ('horror').
Supporting Evidence	The existence of the root in other Indo-European languages, particularly Slavic.	The semantics of dialectal forms: կրկուի կյալ [k'erkup' kyal] ('to fall silent'), կուրկուի ինիլ [kurkup' inil] ('to huddle up'), and use for 'rain holding off' all support the core idea of "closing" or "sealing."
Weaknesses / Unexplained Issues	Does not adequately explain the presence of /r/ in the related verb կուրիլ [kuprēl]. An alternative theory suggests կուրիլ [kuprēl] could be linked to կուրիլ [kuprē] ('tar, pitch'), which would invalidate its connection to this root.	Presents a more internally consistent and complete explanation, with strong parallels within Armenian word formation.

As the comparative table illustrates, the second hypothesis—which posits an origin from reduplication followed by phonetic dissimilation—offers a more robust and internally consistent explanation. It is supported by well-attested phonetic processes within Armenian (e.g., թաղթաղ [t'art'ap'], սարսափ [sarsap']) and demonstrates strong semantic coherence with the word's actual usage in the Goris dialect. In contrast, the first hypothesis, while providing plausible external cognates, struggles to account for the internal phonological details of related dialectal forms and the specific mechanism for the compound's formation.

IV. ANALYSIS AND OBSERVATIONS

The examination of the examples provided in Table 2 and the general lexicon allows for several important conclusions to be drawn regarding the relationship between the Goris and Artsakh dialects, particularly concerning genuine compound words.

A. Shared Lexical Fund and Structural Similarities

Both dialects possess a significant common lexical fund of genuine compound words, which strongly attests to their close historical and linguistic ties. The underlying word-formational models for these compounds are almost identical, suggesting a shared evolutionary path for their compounding mechanisms. This consistency in structural patterns reinforces their classification within the same dialectal sub-branch.

B. Phonetic Correspondences and Divergences

While sharing structural commonalities, the primary distinctions between the two dialects manifest at the phonetic level, particularly concerning vowel usage and transformation.

(a). The Role of the Vowel 'ĕ' (ə)

The vowel 'ĕ' (ə) is highly active in both dialects, yet its manifestation differs. In the Goris dialect, it predominantly serves as a phonetically altered variant of the connecting vowel *a* (e.g., ազկըծուսը [azk'etsup'ĕ] 'the whole family, clan, lineage'), directly participating in the formation of compounds. In contrast, in the Artsakh dialect, it frequently appears as an initial epenthetic addition (e.g., ըզըզամըր [ĕzĕzam'ĕr] 'the beloved has died'), serving a more phonological rather than purely morphological function at the beginning of words. This subtle difference highlights distinct phonological processes at play.

(b). Vowel Alternation and Instability

A notable feature is the frequent instability and alternation of vowels. For instance, the 'ĕ' in Goris (քէթկալ [k'ĕt'kal] 'nose clip / a type of veil') can correspond to the 'ĕ' (ə) in Artsakh (քըթկալ [k'ĕt'kal]). Similarly, Goris's 'i' (as in դիբալ [dibālūs] 'at dawn') might correspond to Artsakh's 'ĕ' (ə) (as in դըբըլ [dĕbĕlūs]). These alternations suggest dynamic phonetic shifts and localized pronunciation preferences within an otherwise common lexical base. Such variations provide valuable insights into the microevolution of vowel systems within closely related dialects.

C. Preservation of Archaic Forms

The pair *պա՛ղպաղոյն* [*pāts'pits'ōru*] (Goris) versus *պէ՛տապէ՛տ* [*pēts'apēts*] (Artsakh) warrants particular attention. In the Goris dialect, the ancient instrumental case ending *'-ōru*' (cf. Classical Armenian *zōr-u*) has been preserved, representing an exceptional archaism. This structure, which links the compound to older morphological patterns, indicates a more conservative linguistic retention in Goris. Conversely, the Artsakh dialect employs the more modern principle of reduplication to convey emphasis or iteration. This stark contrast suggests that the Goris dialect, in certain instances, can be more conservative and retain older linguistic features that have been lost or transformed in other related dialects. This finding is crucial for understanding the historical development and divergence of Armenian dialects.

V. CONCLUSIONS

The structural-semantic examination of genuine compound words in the Goris dialect allows for the following main conclusions:

Genuine compound words in the Goris dialect, being synthetic-type formations, on the one hand, follow the general word-formational regularities of Modern Armenian. They exhibit types with and without connecting vowels, as well as subordinate relationships between their components. The most productive word-formational models are the "noun+noun," "adjective+noun," and "noun+verbal root" types.

On the other hand, the dialect demonstrates a number of unique and exceptional features. Key among these is the phonetic behavior of the connecting vowel *a*, particularly its consistent transformation to *'ē'* (ə) in unstressed positions (e.g., *սրտըկտուր* [*sērt'ēkutur*] 'heartbroken', *քիսկըրիկին* [*ēkhpērēkin*] 'brother's wife'). This phenomenon, contrasting with the canonical *e* connecting vowel in the literary language, imparts a distinctive phonetic color to the dialect and serves as one of its most important distinguishing characteristics.

The comparative analysis of the Goris and Artsakh dialects confirmed their genetic proximity, which is manifested through shared vocabulary and similar word-formational models. Simultaneously, subtle phonetic differences were revealed (e.g., distinct manifestations of the vowel *'ē'* (ə), vowel alternations) and, in some cases, the more conservative nature of the Goris dialect (e.g., the preservation of archaic endings in the word *պա՛ղպաղոյն* [*pāts'pits'ōru*]).

Finally, this work once again confirms that dialects are not corrupted versions of the literary language, but rather independent, rich, and dynamic systems. The study of the word-formational richness of the Goris dialect is important not only from the perspective of the history of the Armenian language but also for preserving its current vitality and diversity, echoing the broader calls for comprehensive dialectal research and preservation efforts (Vaux, 2024).

This work can serve as a foundation for further investigations. Promising areas for future study could include a comparative analysis of non-genuine and fossilized compound words in the Goris dialect, as well as an examination of the usage of these compounds in the speech of contemporary dialect speakers to ascertain their degree of vitality.

APPENDIX

ARMENIAN transliteration table

Uppercase	Lowercase	Uppercase	Lowercase	Uppercase	Lowercase	Uppercase	Lowercase
Ա- A	ա- a	Լ- L	լ- l	Շ- Sh	շ- sh	Բ- W	ւ- w
Բ- B	բ- b	Խ- Kh	խ- kh	Ո- O	ո- o	Ու- U	ու- u
Գ- G	գ- g	Ծ- Ts	ծ- ts	Չ- Ch'	չ- ch'	Փ- P'	փ- p'
Դ- D	դ- d	Կ- K	կ- k	Պ- P	պ- p	Ք- K'	ք- k'
Ե- E	ե- e	Հ- H	հ- h	Ջ- J	ջ- j	Էւ- Ew	եւ- ew
Զ- Z	զ- z	Ջ- Dz	ձ- dz	Ր- R	ր- r	Եվ- Ev	եվ- ev
Է- Ē	է- ē	Դ- Gh	դ- gh	Ս- S	ս- s	Օ- O	օ- o
Ը- Ē	ը- ē	Ճ- Ch	ճ- ch	Վ- V	վ- v	Ֆ- F	ֆ- f
Թ- T'	թ- t'	Մ- M	մ- m	Տ- T	տ- t	Ու- Ü	ու- ü
Ժ- Zh	ժ- zh	Յ- Y	յ- y	Ր- R	ր- r	Ա- Ä	ա- ä
Ի- I	ի- i	Ն- N	ն- n	Յ- Ts'	յ- ts'	Օ- Ö	օ- ö

ACKNOWLEDGEMENTS

The research was supported by the Higher Education and Science Committee of MESCS RA (Research project № 25RG-6B137). The work was carried out within the framework of the topic "Folklore, Dialect, Toponyms, and Family Names of the Goris Region".

REFERENCES

- [1] Acharyan, H. (1971). *Hayeren Armatakan Bararan* [Armenian Root Dictionary] (Vol. 1). Yerevani hamalsarani hratarakhch'ut'yun.
- [2] Aghayan, E. (1984). *Ėndhanur yev haykakan baragitut'yun* [General and Armenian lexicology]. Yerevan State University Press.
- [3] Austin, P. K., & Sallabank, J. (Eds.). (2011). *The Cambridge Handbook of Endangered Languages*. Cambridge University Press.
- [4] Booij, G. (2018). *The Grammar of Words: An Introduction to Linguistic Morphology* (4th ed.). Oxford University Press.

- [5] Chambers, J. K., Trudgill, P., & Schilling, N. (2021). *The Handbook of Language Variation and Change* (3rd ed.). Wiley Blackwell.
- [6] Eloyan, S. (1970). Hodakapov yev an'hodakap bardut'yunnerē zhamanakakits' hayerenum [Compound words with and without connecting vowels in modern Armenian]. *Herald of Social Sciences*, 7, 77–85.
- [7] Greppin, J. A. C., & Khachaturian, A. A. (1986). *A Handbook of Armenian Dialectology*. Caravan Books.
- [8] Hovsepyan, L. (2016). *Grabari barakazmut'yan verlutsakan tipē* [The analytical type of Grabar word formation]. Asoghik Publishing House.
- [9] Jahukyan, G. (1972). *Hay barbaragitut'yan neratsut'yun* [Introduction to Armenian dialectology]. Academy of Sciences of Armenian SSR Press.
- [10] Jahukyan, G. (1987). *Hayots' lezvi patmut'yun, Nakhagrayin zhamanakashrjan* [History of the Armenian language, Pre-literary period]. Language Institute of H. Acharyan, NAS RA.
- [11] Jahukyan, G. B. (1989). *Zhamanakakits' hayots' lezvi imastabanut'yun yev barakazmut'yun* [Semantics and Word Formation of the Modern Armenian Language]. Armenian SSR Academy of Sciences Press.
- [12] Jahukyan, G. (2010). *Hayeren Stugabanakan Bararan* [Armenian Etymological Dictionary]. Asoghik.
- [13] Katvalyan, V. L. (2014). *Hay barbaragitut'yan harc'er* [Problems of Armenian Dialectology]. Yerevan.
- [14] Khachatryan, A. (2020). Hamadrakan yev verlutsakan barakazmakan kaghaparatiperē Barsegh Kesaratsu "Girk' pahots'" yerki grabar t'argmanut'yan mej [Synthetic and analytical word-formational patterns in the Grabar translation of Basil of Caesarea's "Book of Vaults"]. *Herald of Armenian Studies*, 1, 218–228.
- [15] Kumunts, M. (2019). Barak'nnakan-stugabanakan mi ditarkum: Germast – gilimastri – getnakhndzor [A lexicological-etymological note: Germast – gilimastri – getnakhndzor]. *Language and Linguistics*, 1, 65–71.
- [16] Kumunts, M. (2022). Gorisi barbarē hayereni barbarneri bazmahatkanish dasakargman hamatek'stum [The Goris dialect in the context of multi-feature classification of Armenian dialects]. *Proceedings of the "Karabakh-Shamakhi or extreme northeastern interdialect group" conference*, 252–269.
- [17] Kumunts, M., Margaryan, I., Miqayelyan, S., Khachatryan, V., Gevorgyan, G., Khachatryan, H., Grigoryan, S., & Nersisyan, L. (2025). Indo-European Origins of Syunik-Artsakh Agricultural Terms. *Forum for Linguistic Studies*, 7(5), 535–548. <https://doi.org/10.30564/fls.v7i5.9112>
- [18] Margaryan, A. (1976). *Gorisi barbarē* [The Goris dialect]. Yerevan State University Press.
- [19] Margaryan, A. S. (1990). *Zhamanakakits' hayots' lezu* [Modern Armenian language (word formation, phraseology, lexicography)]. YSU Press.
- [20] Mkrtchyan, R. (1972). Bard bāneri dasakargumē hayerenum [Classification of compound words in Armenian]. *Herald of Social Sciences*, 3, 100–106.
- [21] Mkrtchyan, R. (1977). Bard bāneri dasakargumē yev nrants' baghadrich'neri imastayin p'okhharaberut'yunē zhamanakakits' hayerenum [Classification of compound words and semantic relationships of their components in modern Armenian]. *Issues of Language and Style*, 4, 77–158.
- [22] Pokorny, J. (1959). *Indogermanisches etymologisches Wörterbuch*. Francke Verlag.
- [23] Sargsyan, A. (Ed.). (2008). *Hayots' lezvi barbarayin bararan* [Dialectal dictionary of the Armenian language] (Vol. 5). NAS RA "Gitutyun" Publishing House.
- [24] Sargsyan, A. (2013). *Gharabaghi barbari bararan* [Dictionary of the Karabakh dialect]. Edit Print.
- [25] Štekauer, P., & Lieber, R. (Eds.). (2005). *Handbook of Word-Formation*. Springer.
- [26] Torosyan, L. (2018). *Ėndardzak bararan germanerēnē-hayerēn* [Extensive German-Armenian dictionary]. Vienna.
- [27] Vaux, B. (2008). *Vowel harmony in the Armenian dialect of Goris* [Unpublished manuscript]. Cambridge University. Retrieved from <https://lingbuzz.net/lingbuzz/007211>
- [28] Vaux, B. (2024). Current state of Armenian dialectology. In H. Dolatian (Ed.), *Adjarian's Armenian dialectology (1911): Translation and commentary* (pp. 67–130). Language Science Press. <https://doi.org/10.5281/zenodo.14008766>
- [29] Vaux, B., & Suhairi, A. A. (2024). Vowel Harmony in Armenian. In H. van der Hulst & N. A. Ritter (Eds.), *The Oxford Handbook of Vowel Harmony* (pp. 905–917). Oxford Academic. <https://doi.org/10.1093/oxfordhb/9780198826804.013.72>
- [30] Weitenberg, J. J. S. (2002). Aspects of Armenian dialectology. In J. Berns & J. van Marle (Eds.), *Present-day dialectology: Problems and findings* (pp. 141–158). De Gruyter Mouton. <https://doi.org/10.1515/9783110904765.141>

Lusine R. Nersisyan was born in Goris, Armenia. She received her PhD in Philological Sciences from the specialized council of the Institute of Language named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

She is currently the Head of the Chair of Philology and Historical Jurisprudence at Goris State University and a Researcher and Editor at the Syunik Center for Armenological Studies. Her primary research interests encompass philology, historical jurisprudence, and dialectology.

Prof. Nersisyan holds the academic title of Associate Professor. E-mail: lusnersraz@gmail.com (ORCID iD: <https://orcid.org/0000-0001-6025-1141>).

Mher S. Kumunts was born in Goris, Armenia, on September 1, 1972. He graduated from the Faculty of Oriental Studies at Yerevan State University, Yerevan, Armenia, in 1997. He received his PhD in Philological Sciences in 2004 from the H. Acharyan Institute of Language of the National Academy of Sciences of the Republic of Armenia (NAS RA).

He is currently a Researcher at the Dialectology Department of the H. Acharyan Institute of Language (NAS RA) and holds the position of Associate Professor at the Chair of Philology and Historical Jurisprudence at Goris State University. He is also the Chairman of the Syunik Center for Armenological Studies and the founder-director of the website <https://www.syunik.net/>. His primary research interests encompass Armenian dialectology, the history of the Armenian language, etymology, and folklore.

Prof. Kumunts holds the academic title of Associate Professor. E-mail: mherkumunts@gmail.com (ORCID iD: <https://orcid.org/0000-0001-5496-3783>).

Ani A. Ghazaryan was born in Goris, Armenia. She is currently affiliated with the Faculty of Armenian Language and Literature at Goris State University.

She is engaged in studies and research within the field of Armenian linguistics. Her research interests focus on the nuances of the Armenian language and literature.

Ms. Ghazaryan welcomes academic correspondence. E-mail: ani.ghazaryan.05@mail.ru (ORCID iD: <https://orcid.org/0009-0008-3009-2117>).

Manuk H. Falakyan was born in Yerevan, Armenia. He is a PhD in Philological Sciences, specializing in the Armenian language, at the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

He works as a Researcher in the Western Armenian Language Department of the Language Institute and is also a Lecturer at Yerevan State University. His current research focuses on the historical development and modern usage of the Western Armenian language.

Mr. Falakyan can be reached for professional collaboration. E-mail: manukfeleqyan@gmail.com (ORCID iD: <https://orcid.org/0009-0008-3510-6426>).

Armene A. Mkrtumyan was born in Yerevan, Armenia. She earned her PhD in Philological Sciences from the Language Institute named after H. Acharyan, National Academy of Sciences of the Republic of Armenia.

She is a Junior Researcher at the Department of Lexicography at the Language Institute and also an Associate Professor at the Department of Armenian and Foreign Languages of the Armenian National Agrarian University. Her research interests include lexicography and modern Armenian.

Prof. Mkrtumyan holds the academic title of Associate Professor. E-mail: arminemkrtumyan7@gmail.com (ORCID iD: <https://orcid.org/0009-0005-0949-6027>).

Leonid G. Telyan was born in Yerevan, Armenia. He holds a PhD in Philological Sciences from the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

He is a Senior Researcher in the Department of the Modern Armenian Language at the Language Institute. His research interests encompass syntax and semantics in modern Armenian.

Prof. Telyan holds the academic title of Associate Professor. E-mail: leonidtelyan1952@gmail.com (ORCID iD: <https://orcid.org/0009-0001-5296-3365>).

Anahit S. Galstyan was born in Yerevan, Armenia. She received her PhD in Philological Sciences from the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

She serves as a Senior Researcher at the Department of Lexicography at the Language Institute named after H. Acharyan. Her research interests focus on lexicography and Armenian linguistics.

Dr. Galstyan can be contacted for academic collaboration. E-mail: galstyananahit876@gmail.com (ORCID iD: <https://orcid.org/0009-0009-8231-5471>).

Susanna K. Tioyan was born in Yerevan, Armenia. She received her PhD in Philological Sciences from the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

She works as a Senior Researcher in the Western Armenian Language Department of the Language Institute. Her research focuses on the Western Armenian language structure and development.

Dr. Tioyan holds the academic title of Associate Professor. E-mail: susannatioyan@gmail.com (ORCID iD: <https://orcid.org/0009-0001-5986-8854>).

Harout K. Marachlian was born in Armenia. He is a PhD Candidate in Sociolinguistics at the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

He is a Junior Researcher in the Western Armenian Language Department of the Language Institute and serves as an adjunct lecturer of Armenian language and literature at the American University of Armenia. His research primarily deals with language preservation methods and strategies.

Mr. Marachlian welcomes professional correspondence. E-mail: harout.k.marashlian@gmail.com (ORCID iD: <https://orcid.org/0009-0009-6797-8947>).

Tereza S. Shahverdyan was born in the village of Aygehat, Alaverdi region, Armenia. She earned the degree of Doctor of Philological Sciences at the Hrachya Acharyan Institute of Language of the National Academy of Sciences of the Republic of Armenia.

She is currently a professor and lecturer at the Chair of Armenian Language and Literature at Vanadzor State University. Her main research interests include Classical Armenian (Grabar), the history of the Armenian language, and historical stylistics. She is the Editor-in-Chief of the Scientific Collection of Works “Teghekagir” of Vanadzor State University.

Prof. Shahverdyan welcomes academic inquiries. E-mail: tereza.shahverdyan@mail.ru (ORCID iD: <https://orcid.org/0009-0007-6713-8994>).

Anna H. Abajyan was born in Armavir (formerly Hoktemberyan), Armenia. She graduated from Yerevan State University. She received her PhD in Philological Sciences in 1992 and her Doctor of Philological Sciences degree in 2006 from the H. Acharyan Institute of Language of the National Academy of Sciences of the Republic of Armenia.

She is currently a Professor at the Faculty of Armenian Philology at Yerevan State University and holds the position of Head of the Chair of History of Armenian Language and General Linguistics. Her research interests include the history of the Armenian language, historical morphology, semantics, word formation, and the theory of language.

Prof. Abajyan holds the academic title of Professor. E-mail: abajyan.anna@ysu.am (ORCID iD: <https://orcid.org/0000-0003-3728-0862>).

Manana G. Dalalyan was born in Vanadzor, Armenia. She received her PhD in Linguistics from Yerevan State University.

She is an Assistant Professor at the English Philology Chair at Yerevan State University, Chair of Foreign Languages and Methodology at ASPU, and Chair of Applied Linguistics of EUA. She is currently the secretary of the scientific council in the Faculty of European Languages and Communication at YSU. Her primary research fields include Lexicology, Stylistics, Text Linguistics, and Pragmatics.

Dr. Dalalyan is available for professional correspondence. E-mail: mananadalalyan@yahoo.com (ORCID iD: <https://orcid.org/0009-0007-8038-0953>).

Vanuhi M. Baghramyan was born in Nor Kharberd, Armenia. She is a PhD Candidate in General and Comparative Linguistics at the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

She is a Junior Researcher in the Lexicography Department of the Language Institute and works as a teacher at the Basic School-College of Khachatur Abovyan, Armenian State Pedagogical University (ASPU). Her research interests include general and comparative linguistics, lexicography, and language history.

Ms. Baghramyan can be contacted for academic inquiries. E-mail: vanuhi.baghramyan@gmail.com (ORCID iD: <https://orcid.org/0009-0003-5821-9574>).

Karine L. Arakelyan was born in Yerevan, Armenia. She holds a PhD in Philological Sciences from the Language Institute named after Hrachya Acharyan of the National Academy of Sciences of the Republic of Armenia.

She works as a Senior Researcher in the Modern Armenian Language Department of the Language Institute. Her research primarily deals with the modern Armenian language structure and development.

Dr. Arakelyan's contact details are listed for professional correspondence. E-mail: karinearakelyan079@gmail.com (ORCID iD: <https://orcid.org/0009-0001-2073-7021>).

Modelling Cognitive and Metacognitive Strategies and Their Relationships to Reading Comprehension Test Performance of Thai High School EFL Learners

Panassanan Kitichaidateanan

Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand

Nithipong Yothachai

Faculty of Education and Human Development, Chaiphum Rajabhat University, Thailand

Apisak Sukying*

Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand

Abstract—Cognitive and metacognitive strategies are crucial for enabling reading comprehension and controlling language processing. This study explored the role of these strategies and their contribution to English reading comprehension test performance, as well as the extent of their relationships in reading comprehension outcomes, using Structural Equation Modelling (SEM). A total of 685 Thai high school students were assessed on a standardised reading comprehension test and a six-point Likert-scale questionnaire designed to measure strategy use. The findings indicated that cognitive and metacognitive strategies were positively related to reading test performance. Cognitive strategies, especially those of retrieval and comprehension, served as basic processes underpinning decoding, integration, and the construction of meaning from text. However, metacognitive strategies exerted significant regulatory influence, with direct effects on reading comprehension and indirect effects via cognitive strategy use. The relationship between retrieval and comprehension also emphasised their interrelated roles as components of strategic reading. Both types of strategy contributed significantly to reading comprehension, but only through language knowledge, with the learning processing model showing incremental improvement in strategy use and apparent individual differences as proficiency grew. These results point to the pedagogical importance of developing strategic awareness and self-regulated learning, as well as providing more explicit strategy instruction, and indicate that future research using longitudinal or mixed-methods designs could provide richer information about how strategies are used differently across contexts, age groups, and more advanced proficiency levels.

Index Terms—metacognitive strategies, cognitive strategies, reading comprehension, Thai high school students, SEM

I. INTRODUCTION

Reading is one of the four major language skills and is central to students' comprehensive academic improvement and language learning. In addition to the reading passages, including information that is related to other subjects, it is also essential for learners in English as a Foreign Language (EFL) contexts to read general, non-academic texts in terms of learners' academic success and development of other language skills such as writing, speaking, listening, and vocabulary advancement because their exposure to authentic English input at school may be quite limited. EFL learners have difficulties understanding written material, making inferences, and generalising information from text to everyday situations. A lack of vocabulary, grammatical knowledge, and background information often compounds these difficulties.

This study investigates different reading strategies that have been shown to facilitate comprehension. Language testing research has produced a rich profile of factors affecting reading comprehension, including linguistic knowledge, processing abilities, and strategic behaviour (Grabe, 2009; Grabe & Stoller, 2020; Kitichaidateanan & Sukying, 2025; Phakiti, 2007). Of these, reading strategies learners consciously direct behaviour that they enact to facilitate the understanding of what they read have been demonstrated to be important (particularly for L2 readers of varying proficiency and learning profiles). Skilled readers are active and self-directing, modifying their reading strategies according to the difficulty of a text, the purpose of the reading (e.g., scanning content or thoroughly learning about a topic), the nature of the task requirements, and progress toward achieving comprehension.

* Corresponding Author. Email: apisak.s@msu.ac.th

Reading strategies are typically broken down into two main categories: cognitive and metacognitive strategies. Cognitive strategies are the mental processes used to make sense of what we read (Kitichaidateanan & Sukying, 2025; O'Malley & Chamot, 1990; Oxford, 2011; Phakiti, 2003, 2006; Sari, 2016; Zhang, 2018; Zhang & Guo, 2020). These are, for example, prediction, summarisation, inference, translation, and the integration of new information with old. Metacognitive strategies, however, are higher-order control processes that empower learners to plan, monitor, and evaluate their own cognitive strategies (Cai et al., 2022; Kitichaidateanan & Sukying, 2025; Noipa & Phusawisot, 2025; Oxford, 2017; Phakiti, 2003, 2006; Zhang, 2018; Zhang & Guo, 2020). Readers who are increasingly aware of their comprehension through metacognitive control can more effectively identify breakdowns and adapt their approach as needed. While metacognitive processing systems may initially seem to retard reading for meaning, they will ultimately lead most learners toward a more effective and independent mode of reading.

Although extensive international research has been conducted, few studies have examined the interaction between cognitive and metacognitive strategies in relation to the reading performance of Thai high school EFL learners. Although earlier studies (Kitichaidateanan & Sukying, 2025; Noipa & Phusawisot, 2025; Phakiti, 2003, 2006; Sukying, 2021) have found positive correlations between strategy use and reading comprehension, ongoing research is needed to disambiguate how these strategies might work together within Bachman and Palmer's (2010) strategic competence construct. To fill this gap, the current study seeks to model the mediating roles of cognitive and metacognitive strategies and their implications for test performance in a reading comprehension test, using Structural Equation Modelling (SEM). The following questions guide this study.

1. What are the relationships between cognitive and metacognitive strategies and reading test performance?
2. To what extent do cognitive and metacognitive strategies affect reading comprehension test performance?

II. LITERATURE REVIEW

A. *Communicative Language Ability*

Communicative language ability (CLA) is a multidimensional, dynamic, and context-dependent construct that pervades the successful use of a second language (L2) in real communicative situations (Bachman & Palmer, 2010; Cai & Zhao, 2022; Purpura, 1999). It is not developed or performed in a social and cultural vacuum, where its development would also be complicated (McNamara, 1996). Over the past decades, applied linguistics and language testing scholars have advanced several important models to conceptualise CLA (Chomsky, 1965) and the ways in which linguistic knowledge interfaces with contextual needs and communicative intentions.

Among the most influential and widely referenced models is the taxonomy by Bachman and Palmer (1996), which classifies CLA into two linked components: language knowledge and strategic competence. Language competence is the knowledge of language held in memory, while strategic competence is the cognitive tools that help program and organise, monitor, and evaluate what has been produced. In their 2010 revision, Bachman and Palmer developed the model further to include language processing capacities/memory functions, but also (1) cognitive methods of knowing, (2) topical learning and knowledge, as well as (3) affective schemata/traits and social attributes that affect test performance and external uses of language.

One of the critical questions in language testing is to what degree these tasks tap into learners' underlying language ability and how well scores validly represent that ability (Bachman, 2000; Purpura, 2014). Within this framework, strategic competence, as operationalised in metacognitive regulation, is central to the management of language processing. Thus, the current study uses Bachman and Palmer's (2010) model to examine the role of metacognitive knowledge in its interaction with L2 proficiency and reading test performance.

B. *Strategy Use*

Strategic processing is a central part of strategic competence in human information processing and is closely related to metacognitive regulation. In language learning, strategic competence is typically implemented through a set of metacognitive strategies that plan, monitor, and evaluate cognitive operations. During active processing in these higher-order processes, learners monitor and control the lower-level cognitive activities of information organisation and integration, thereby increasing their ability to process complex language tasks. One crucial difference between strategic and automatic processing is awareness; whereas automatic processes occur quickly, with little attentional demand, strategic processing is conscious, goal-driven, and intentional.

The dynamic interplay between automatic and strategic processing is evident in L2 reading comprehension. Readers rely on both lower- and higher-level processes to build meaning, alternating flexibly between unconscious and conscious processing as the demands of activities change (Alderson, 2000; Phakiti, 2007; Song & Cheng, 2006). Automatic processes support rapid retrieval of word forms and grammatical structures, as well as the development of text coherence. Heterarchical processing is mandatory when learners encounter new words, complex grammar, or comprehension breakdowns. Strategic processes, for example, inferencing, problem-solving, monitoring comprehension of text, and repairing misunderstandings, are influential contributors to successful (and long-term) reading development.

Phakiti (2007) defines strategy as one of the essential components of the control processing system, governing, controlling, and regulating cognitive events in language use. Strategies in this sense are not subconscious, haphazard, or

intuitive, but rather partly conscious, intentional, and voluntary activities used to facilitate communication and learning. As a result, the strategic use of the role is highly regarded as an indicator of successful L2 performance.

Cognitive strategies are essential tools that enable learners to consciously control, manipulate, and build meaning during L2 learning (Kasimi, 2012; Song & Cheng, 2006). These strategies assist with reading comprehension, storing and recalling information—all key components of successful L2 development (Macaro, 2006; Phakiti, 2003; Song & Cheng, 2006). According to the revised language use construct proposed by Bachman and Palmer (2010), cognitive strategies are understood as concrete, observable behaviours used during the performance of a language task. Although classified as “peripheral” rather than “focal,” they are, however, essential for accounting for variance in language test scores and for linking test achievement to real-world proficiency.

The conceptual framework of cognitive strategies is cognitive learning theory, which holds that learning consists of actively managing (manipulating, organising, transforming) inputs to improve comprehension and memory. In this view, cognitive strategies are conscious mental operations that allow learners to process language information more efficiently (Anderson, 2008; Anderson & Pearson, 1984; O'Malley & Chamot, 1990; Kitichaidateanan & Sukying, 2025; Macaro, 2006). Wenden (1991) and Purpura (1999) also describe cognitive strategies as mental operations and observable behaviours that assist in processing, storing, and retrieving linguistic information. Recent empirical studies have also provided additional evidence of their importance for understanding what is read: students who actively activate relevant schema, attend to grammatical signals, make inferences from context, and monitor their comprehension often achieve better comprehension outcomes (Phakiti, 2016; Purpura, 2014; Sukying, 2021). Collectively, these theoretical and empirical findings position cognitive strategies as core processes by which learners make sense of textual information, connect new information to existing cognitive structures (i.e., mental models), and construct coherent meaning during reading.

Metacognitive strategies operate at a superordinate/executive level and concern the planning, monitoring, and management of learning operations. They are global processes that guide and regulate lower-level cognitive processes (Sheorey & Mokhtari, 2001). To this date, some controversy exists among researchers in the field regarding how specific strategies can be categorised, as predicting meaning or using textual cues, for example, is placed in either the cognitive or the metacognitive strategy category (Phakiti, 2003). Cognitive strategies mainly facilitate processing information, and metacognitive strategies are responsible for controlling and utilising this processing to comprehend (Anderson, 2009; Kitichaidateanan & Sukying, 2025; Purpura, 2014).

A second important consideration is whether strategies are trait- or state-based. Previous studies treated strategies as stable, trait-like preferences, whereas more recent research (Oxford, 2017; Phakiti, 2008; Sukying, 2021) has shown that strategy use is dynamic, situational, and task-dependent. In line with this view, the present study takes a state-based perspective by investigating learners' actual deployment of cognitive and metacognitive strategies during reading test tasks.

By framing metacognitive strategies as dynamic, context-dependent regulatory procedures, the present study aligns with contemporary theoretical and empirical trends, emphasising the adaptive nature of strategy use and its pivotal role in L2 reading success.

C. EFL Reading

Reading in a second language (L2), especially in EFL, is a complex and multifaceted cognitive activity influenced by various linguistic and non-linguistic factors. It demands decoding, as well as complete control over syntax, semantics, discourse structures, and strategic activity based on reading processes. Successful EFL reading is also contingent on the building of coherent mental representations of text, triggering the activation of both topical and background knowledge, and employing metacognitive control to oversee and regulate comprehension.

Many factors affect EFL reading ability. Factors related to the reader include the influence of the first language, English experience, L2 reading proficiency, motivation, and awareness of reading strategies. Text-based variables, such as topic familiarity, text difficulty, genre, and the addition of multimodal support (e.g., images), also interact with comprehension outcomes. Contextual factors, such as what type of text readers would usually encounter (e.g., newspaper vs. academic), could scaffold or impede comprehension depending on readers' familiarity with it (Alderson, 2000; Phakiti, 2007). The cognitive processes involved in L1 and L2 reading are similar in many ways (Grabe & Stoller, 2020; Song & Cheng, 2006); however, due to language limitations, increased processing load, and lower automaticity, EFL reading is often slower and less fluent. Research evidence has consistently indicated that reading in EFL involves an interactive, dynamic process influenced by learners' interlanguage state, learner differences, and the broader social context of instruction (Alderson, 2000; Kitichaidateanan & Sukying, 2025; Koda, 2005; Noipa & Phusawisot, 2025).

In light of these complexities, an examination of EFL reading processes is a crucial requirement if we are to determine what readers do when they read and how test tasks tap into this cognitive and metacognitive activity. Thus, in response to these gaps, this study, using structural equation modelling (SEM), investigates how relationships exist among cognitive and metacognitive strategy use, strategic knowledge, regulatory processes, and their performances on the English reading test for a group of high-school EFL learners in Thailand.

III. RESEARCH METHODOLOGY

A. Participants and Setting

The participants in this study were 685 Grade 12 students aged 17-18 from a public high school in northeastern Thailand. The participants' English skills ranged from advanced beginners to upper-intermediate, and their families came from various socioeconomic and occupational backgrounds. All participants had completed at least 10 years of EFL courses, as required by the Office of the Basic Education Commission (Ministry of Education of Thailand, 2008). They attended four 50-minute English classes with EFL teachers and one class with native English speakers per week, with class sizes ranging from 20 to 45 students. The students were familiar with effective reading strategies and continued their English studies at a higher academic level. Convenience sampling was used to select participants for this study because it is efficient for quickly gathering basic information (Dörnyei, 2007; Sekaran & Bougie, 2016).

B. Research Instruments

(a). Reading Comprehension Test

The Reading Comprehension Test (RCT) is a standardised measure designed to assess the reading comprehension abilities of Thai secondary school learners. The test was constructed according to good language testing practices, and consultation with five expert EFL teachers (≥ 10 years' experience) had taken place. The RCT is composed of 60 items on two subtests: (a) a 20-item rational cloze designed to evaluate lexico-grammatical reading ability and (b) a comprehension section assessing prominent idea recognition, supporting detail recognition, and inferencing.

A pilot study was conducted on a sample of 168 non-sample students, establishing content validity through Item-Objective Congruence (IOC). Items that did not meet the 0.60 criterion were revised or eliminated to maintain high instructional alignment. The item analysis revealed an adequate difficulty and discrimination, with all items higher than the minimum of .20 in the case of the discrimination index ($r = .51$). Difficulties were calibrated to ensure equal distribution of 25% difficult, 50% moderate and 25% easy items.

To determine whether all instrument subscales showed normal distributions, mean scores and skewness values were analysed. Cronbach's alpha for the reliability of the scale was $\alpha = .94$, indicating excellent internal consistency. The test, constructed in gap-fill and multiple-choice formats within an hour, has been systematically developed to form a valid, reliable, and psychometrically robust measure of EFL reading comprehension for Thai high school students.

(b). A Cognitive and Metacognitive Strategy Use Questionnaire

The questionnaire was adapted from Phakiti's (2007) validated instrument and incorporated items from established research on language learning strategies (Sukying, 2021), reading strategies (Mokhtari & Reichard, 2002; Phakiti, 2003, 2008), and test-taking strategies (Anderson, 1991; Anderson et al., 1991; Cohen & Upton, 2006). The theoretical foundation was human information-processing theory, distinguishing between cognitive structures (sensory input, working memory, long-term memory) and functional processes (comprehension, retention, retrieval, and executive control).

To minimise linguistic load, the questionnaire was administered in Thai. Items were rated on a six-point Likert scale (0-5), consistent with Phakiti's (2007) recommendations for capturing learners' strategy use. The instrument contained 51 items across six subscales: comprehending (6), memory (8), retrieval (7), planning (10), monitoring (14), and evaluation (6). Completion time ranged from 10 to 15 minutes.

A pilot study with 168 non-sample students confirmed strong reliability. Cronbach's alpha coefficients for the subscales ranged from .81 to .97, with an overall coefficient of .99, demonstrating excellent internal consistency. These results indicate that the instrument is a psychometrically robust and well-validated measure of learners' cognitive and metacognitive strategy use.

C. Data Collection

The students and their parents provided written consent prior to data collection, after formal approval from the school authorities. All participants completed the RCT during the first week of class. The test lasted one hour and was conducted in a standardised testing classroom. Two experienced English teachers administered each test as proctors to guarantee standardised execution and compliance with the avoidance of using strategies.

The Cognitive and Metacognitive Strategy Use Questionnaire was administered immediately at the end of the Reading Comprehension Test to evaluate participants' strategy use throughout testing. To minimise the influence of quality on data collection, responses from participants who provided an alternative answer to 10 or more consecutive RCT items were excluded. These responses were viewed as non-serious.

D. Data Analysis

Participants' performance and strategy use were initially summarised using descriptive statistics (means and standard deviations), followed by inferential analysis to address the research questions. The application used two analysis targets: (1) to model and validate reading comprehension test performance with the Structural Equation Modelling (SEM) process;

and (2) to detect and treat examinees who do not fit to ensure analytical soundness. Problematic responses were omitted before SEM to reduce measurement error and improve model fit.

The dimensionality of the item groups was verified in the first stage through Confirmatory Factor Analysis (CFA), yielding composite variables with adequate measurement properties. Additionally, we conducted outlier diagnostics to identify observations that could have an undue impact on parameter estimates.

SEM adopted a single-group approach to test the measurement and structural components of the hypothesised model. This procedure comprised model specification, maximum-likelihood parameter estimation, and fit testing based on multiple criteria. These measures led to a thorough examination of the hypothesised relationships between the use of cognitive and metacognitive strategies and EFL reading performance, yielding valid and reliable inferential conclusions.

IV. RESULTS

A. Descriptive Statistics

Table 1 presents descriptive statistics on the use of cognitive and metacognitive strategies reported by Thai high school students in the reading section of a comprehension test. In sum, participants reported a moderate frequency of strategy use on average ($M = 4.13$, $SD = 0.81$). That is, reading comprehension (74.64%) was the most common strategy employed; monitoring (69.86%), memory (69.73%), and retrieval (68.63%) and planning strategies were also frequently used (67.53%). Comparison as an evaluating method was the least frequently used strategy (62.42%). These results suggest that Thai EFL learners do employ practical strategies to read in meaningful ways and to monitor their reading.

Learners reported using overall cognitive strategies ($M = 4.26$, 71%) more frequently than overall metacognitive strategies ($M = 4.00$, 66.6%) across the two main factors. This result indicates that, as a mode of learning, cognitive strategies are more frequent than higher-level control strategies of the processes involved in actual text reading. These findings also provide evidence that Thai high school learners are more balanced in their adoption of both sides, with a cognitive bias toward reading comprehension. Their low use of metacognitive strategies reflects a lack of self-regulation, and a high level of cognitive processing demonstrates the role of direct decoding.

TABLE 1
OVERALL STRATEGY USE OF THAI HIGH SCHOOL LEARNERS

Strategy use	Mean	%	S.D.
Comprehending	4.48	74.64	0.65
Memory	4.18	69.73	0.95
Retrieval	4.12	68.63	1.04
Cognition	4.26	71.00	0.84
Planning	4.05	67.53	0.76
Monitoring	4.19	69.86	1.02
Evaluating	3.75	62.42	0.81
Metacognition	4.00	66.60	0.81
Overall	4.13	68.80	0.81

Note: $N = 685$

Table 2 presents the reading comprehension test (RCT) performance of Thai high school students across all components. The highest average value was also achieved for LexGrRA2 (41.90%), immediately after TxtCOMP (40.85%) and LexGrRA1 (40.67%). Note that although some fluctuations were observed between sections, all mean scores were below half of the maximum score that could be attained, which is indicative of the very challenging nature of both lexico-grammatical and text comprehension tasks for learners.

The average test score at the overall level was 40.99%, suggesting low reading comprehension among Thai high school participants. Though students showed some ability to interpret lexico-grammatical information and derive meanings from texts, they were not proficient enough to understand the readings effectively. These findings point to ongoing difficulties with both word-level processes (e.g., vocabulary and grammatical decoding) and comprehension skills (e.g., inferencing, integrating ideas, and understanding relationships between texts). Overall, the findings highlight the importance of targeted instructional intervention to develop related skills in foundational and integrative reading for these Thai high school readers.

TABLE 2
READING COMPREHENSION TEST PERFORMANCE OF THAI HIGH SCHOOL LEARNERS

Reading comprehension test	Total	Mean	%	S.D.	Min	Max
LexGrRA1	10	4.07	40.67	2.293	0	10
LexGrRA2	10	4.19	41.90	2.483	0	9
TxtCOMP	40	16.34	40.85	6.995	5	32
RCT	60	24.60	40.99	10.536	9	46

Note: LexGrRA = Lexical-Grammatical reading ability, TxtCOMP = Text comprehension ability
LexGrRA1, LexGrRA2 and TxtCOMP = Reading Comprehension Test (RCT)

B. Relationship Between Strategy Use and Reading Comprehension Test Performance

Before testing the proposed structural relationships, individual measurement models were checked for model fit and the possibility of misspecification within the overall latent SEM. Each model was specified and estimated for all observed variables using the maximum likelihood method, applicable to ordinal-scale data and approximately multivariate normal data. The maximum likelihood estimation reduces bias in parameter estimates by using a population-replicating variance-covariance matrix. Goodness-of-fit was tested using chi-square values, standardised residuals, and fitness indices well established in the literature and supported by both theoretical and empirical reasons.

After obtaining good measurement models, the structures were examined. The effect of metacognitive strategies on cognitive strategies was first estimated, and then the impact of cognitive strategies on reading comprehension performance was assessed. The first model (speculated) was not confirmed, and this omission left us wondering what prompted us to study further models. Conditional error correlations were added to the model, following Bentler (2006), to account for correlated measurement errors, thereby improving fit and explanatory power.

A satisfactory fit was obtained with the final SEM model (Figure 1). Although the Chi-square test was statistically significant ($\chi^2 = 99,053.55$, $p < .001$), and, as is common with large samples, the remaining showed a close fit: $\chi^2/df = 2.95$, CFI = .90, TLI = .91, RMSEA = .05 (90% CI = .046-.054), and SRMR = .049. These values met the standard acceptable-to-good fit criteria (Hair et al., 2019). All parameter estimates were significant ($p < .05$), supporting the hypothesised relationships.

Table 3 summarises the standardised path coefficients and R^2 for the exogenous constructs. R^2 was used as an effect-size estimator, representing the proportion of variance explained in the dependent variables. For interpretive sense-making, the observed events were coded descriptively (i.e., PLA1 = plan before reading). In sum, the results offer solid empirical support for an indirect effect of metacognition on reading comprehension through its regulation of cognitive strategy use and testify to the essential role of strategic processing in Thai high school students' comprehension performance.

TABLE 3
STANDARDISED PARAMETER ESTIMATES FOR STRATEGY USE AND READING COMPREHENSION TEST PERFORMANCE

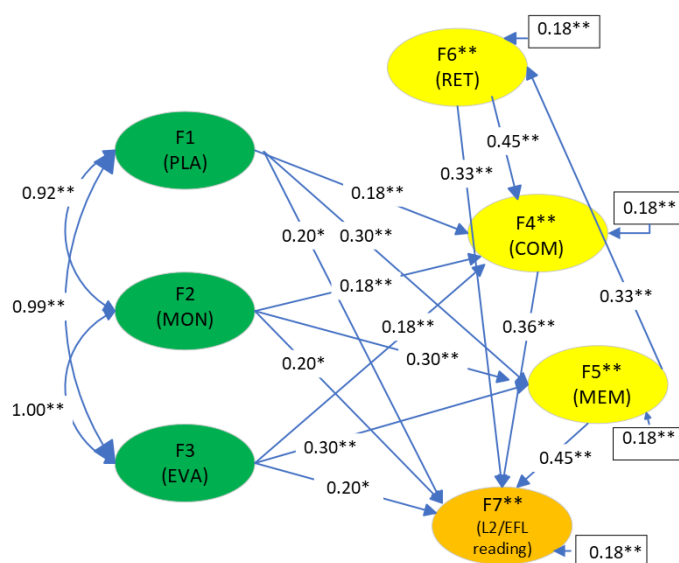
Observed variable		Label	Loading	Error	R ²
Item1	PLA1	plan before reading	0.94*F1	+0.12 E1	0.88
Item2	PLA2	understand the goal of the task	0.84*F1	+0.29 E2	0.71
Item3	PLA3	think about what to achieve	0.83*F1	+0.31 E3	0.69
Item4	PLA4	know what to do and how to do	0.80*F1	+0.35 E4	0.65
Item5	PLA5	realise whether or not plans work well	0.82*F1	+0.33 E5	0.67
Item7	PLA7	identify purposes	0.76*F1	+0.42 E7	0.58
Item9	PLA9	highlight the length and arrangement of a text	0.75*F1	+0.43 E9	0.57
Item11	COM1	look for the main idea in the first sentence	0.99*F4	+0.02 E11	0.98
Item12	COM2	figure out the link to the main ideas	0.99*F4	+0.01 E12	0.99
Item15	COM5	Predict what will happen next	0.97*F4	+0.06 E15	0.94
Item16	COM6	interpret the author's intended message	0.97*F4	+0.05 E16	0.95
Item17	MEM1	use typographic features	0.99*F5	+0.04 E17	0.96
Item18	MEM2	reread the text to understand it better	0.90*F5	+0.20 E18	0.80
Item20	MEM4	try to decipher hidden concepts	0.88*F5	+0.22 E19	0.78
Item25	RET1	use prior knowledge	0.94*F6	+0.04 E25	0.96
Item26	RET2	aware of relevant information	0.95*F6	+0.04 E26	0.96
Item27	RET3	use context clues	0.97*F6	+0.06 E27	0.94
Item28	RET4	use the grammar rules	0.97*F6	+0.07 E28	0.93
Item29	RET5	know root words	0.96*F6	+0.07 E29	0.93
Item30	RET6	infer the information	0.98*F6	+0.04 E30	0.96
Item31	RET7	comprehend from prior knowledge	0.90*F6	+0.19 E31	0.81
Item32	MON1	aware of the time limitations	1.00*F2	+0.01 E32	0.99
Item33	MON2	aware of the amount of reading and tasks	0.84*F2	+0.29 E33	0.71
Item34	MON3	aware of when and where it is confusing	0.88*F2	+0.23 E34	0.77
Item35	MON4	know when you get anxious or uninterested	0.85*F2	+0.28 E35	0.72
Item36	MON5	know when losing attention	0.78*F2	+0.38 E36	0.62
Item37	MON6	double-check, or self-reflect	0.83*F2	+0.32 E37	0.68
Item38	MON7	aware of whether or not you comprehend the text	0.86*F2	+0.25 E38	0.75
Item39	MON8	pay attention to the necessary detail	0.86*F2	+0.25 E39	0.75
Item40	MON9	adapt reading speed	0.88*F2	+0.32 E40	0.68
Item41	MON10	manage the time	0.78*F2	+0.39 E41	0.61
Item42	MON11	correct the mistake effectively	0.81*F2	+0.34 E42	0.66
Item43	MON12	adjust reading speed	0.80*F2	+0.35 E43	0.65
Item44	MON13	adapt the pace of answering	0.80*F2	+0.36 E44	0.64
Item45	MON14	use context clues	0.84*F2	+0.30 E45	0.70
Item46	EVA1	prove comprehension of the task	0.90*F3	+0.19 E46	0.81
Item48	EVA3	evaluate reading strategies	0.53*F3	+0.72 E48	0.28
V52	LexGrRA1	Lexical-Grammatical reading ability1	0.53*F7	+0.72 E52	0.87
V53	LexGrRA2	Lexical-Grammatical reading ability2	0.53*F7	+0.72 E53	0.92
V54	TxtCOMP	Text comprehension ability1	0.53*F7	+0.72 E54	0.90
F4	Com	Comprehending strategies	0.18*F1 +0.18*F2 +0.18*F3 +0.45*F6	+0.06 D1	0.94
F5	Memory	Memory strategies	0.30*F1 +0.20*F2 +0.30*F3	+0.18 D2	0.82
F6	Retrieval	Retrieval strategies	0.20*F1 +0.20*F2 +0.20*F3 +0.33*F5	+0.18 D3	0.82
F7	RCT	Reading strategies	0.36*F4 +0.22*F5 +0.33*F6	+0.22 D4	0.78

In the final SEM model (Figure 1), there were unidirectional relationships among memory, retrieval, and comprehension strategies, which did not support our initial supposition of reciprocal relationships. Memory strategies directly predicted retrieval strategies ($\beta=.33$, $R^2=.11$). They indirectly affected comprehension strategies then the recovery ($\beta=.15$, $R^2=.02$). The retrieval strategies had the strongest effect on comprehending strategies ($\beta=.45$, $R^2=.20$), which positions retrieval as the single most crucial vessel connecting represented (stored) knowledge to meaning making.

The measurement model well represented cognitive strategies. Understanding strategies (e.g., understanding main ideas, connecting information, interpreting the author's purpose) and retrieval strategies (e.g., activating prior knowledge, using grammar rules, and using context clues) exhibited very strong factor loadings (.90-99), explaining 93-96% of the variance. Memory strategies (i.e., rereading, typographic cues) were also relatively strong, accounting for 85% of the variance.

Metacognitive strategies displayed greater variability. Both planning (goal setting, task organisation) and monitoring (time management, attention regulation, self-reflection) explained 71% of the variance, with evaluating strategies the weaker construct (54%). This indicates that learners are constantly planning and monitoring their reading, but do not

evaluate their own performance. These results suggest that comprehension in Thai high school learners is mediated by retrieval strategies that modulate the influence of memory on comprehension. Whereas cognitive strategies were highly reliable, metacognitive regulation, including evaluation, was inconsistently supported, suggesting a potential area for pedagogical development.



Chi-square=99053.55, df=742, p=0.000, TMSEA=0.446

Figure 1. The Hypothesised Model of the Relationships of Strategy Use and Reading Comprehension Test Performance

Note: F1 (PLA) = Planning strategies F2 (MON) = Monitoring strategies,

F3 (EVA) = Evaluating strategies F4 (COM) = Comprehending strategies

F5 (MEM) = Memory strategies F6 (RET) = Retrieval strategies

F7 (L2/EFL reading) = Reading Comprehension Test

V. DISCUSSION

The current study examines the associations among the use of cognitive and metacognitive strategies for reading comprehension among Thai high school students by investigating the predictive power of these strategies on reading comprehension and their interrelations using structural equation modelling (SEM). The results offer a holistic view of the functions of strategies in reading and how they influence learners' test performance. In sum, the findings suggest that strategy use does have some impact, albeit an indirect one, on reading comprehension.

The present study's results indicated that metacognitive and cognitive strategies were significantly positively associated with reading variables. This finding can probably be attributed to the fact that students who practised these strategies more frequently performed at higher levels of reading comprehension. Cognitive processes like inference, summarisation, retrieval, and activation of prior knowledge are dynamic mechanisms that specifically contribute to making sense of what one reads by aiding the reader in decoding information from written texts and making interpretative sense. Metacognitive strategies also mediate performance, as they allow readers to plan their reading, monitor their comprehension, and review their progress, thus controlling cognitive strategy use to maximise efficiency. The relationships between the two types of strategies found are all positive, indicating that comprehension can be explained by a general processing capacity beyond what is explicable in terms of handling linguistic analogies, particularly in terms of knowing relations. Together, these findings underscore the interactional nature of cognitive and metacognitive strategies in promoting comprehension and stress the need to focus on strategic knowledge and self-regulation skills of learners to improve overall reading ability.

The SEM analysis revealed that metacognitive strategies had direct and indirect influences on reading comprehension through cognitive strategy use. Specifically, students who planned their reading, monitored comprehension, and evaluated their progress were more likely to use specific cognitive strategies, such as inferring, to better understand the text. While the direct effects of strategic behavior ($\beta = .58-.86$), although these effects were significant and in favor of their predictive capacities, they were smaller than language knowledge suggesting that strategy use enhances performance but cannot supplant the pivotal role of language mastery (Kitichaidateanan & Sukying, 2025; Lin et al., 2019; Phakiti, 2008; Purpura, 1999; Song, 2005).

The direct and indirect effects of metacognitive strategies on reading comprehension via cognitive strategy use might be explained by the fact that metamemory processes involved in actively controlling, orchestrating, monitoring, and regulating reading are critical, multifaceted categorical variables that determine learners' reading performance. This direct effect is interpreted to mean that learners who plan their reading, monitor what they have read, and evaluate the progress of their understanding are successful in comprehension, irrespective of the cognitive strategies used. Through these

regulatory behaviours, learners can more effectively establish reading goals, recognise when comprehension has broken down, and enact repairs that directly contribute to their understanding of the text.

The mediating effect indicates that metacognitive strategies support and consolidate the application of cognitive strategies. The SEM results showed that students who utilised more metacognitive regulation also tended to employ appropriate cognitive strategies, including inferring meaning, retrieving relevant knowledge, summarising main ideas, and integrating text information with their prior knowledge. Once activated, these cognitive processes provided direct access to learners' understanding. Hence, the relationship between metacognitive strategies and reading is very complex, as they contribute not only to their own regulatory role but also to the more effective utilisation of cognitive strategies.

Taken together, these routes suggest that metacognitive strategies constitute a transtheoretical control system that regulates cognitive processing in reading. Their integrated direct and indirect effects illustrate the complexity of successful comprehension and underscore why learners need explicit instructional support in regulating their strategy use. Thus, building metacognitive awareness and self-regulated learning abilities has the potential to improve reading comprehension by increasing readers' ability to monitor comprehension and to deploy cognitive strategies more effectively.

The structural relationships between cognitive strategies demonstrated a clear hierarchical pattern, in which memory processes had a significant impact on retrieval, and retrieval had an overwhelming effect on comprehension. This hierarchy implies that retrieval is the central cognitive process of reading, showing that good comprehension depends less on simple memorisation and more on the capacity to retrieve, activate, and use stored linguistic and contextual knowledge during real-time language processing. This reading is consistent with Anderson's (2009) conceptual understanding of reading as a form of information processing in which retrieval is the first and necessary process for meaningful comprehension of text. It is also in line with the notion that strategic L2 use and controlled processing are essential for handling reading tasks and enhancing comprehension performance (Kitichaidateanan & Sukying, 2025; Phakiti, 2003, 2008; Zhang et al., 2014). In sum, these findings suggest that although several cognitive strategies contribute to reading comprehension, they are not equivalent; retrieval is the primary mechanism by which memory resources are converted into a coherent understanding of the text.

Metacognitive strategies also showed a coherent internal structure, with the shared variances among planning, monitoring, and evaluation being extreme. Monitoring was also the most robust factor, in line with Oxford's (2017) S²R model, which regards monitoring as a key strategy for governing and modifying success strategies. The evaluative dimension, however, was the weakest, with reiterations of the reported difficulties in self-assessment with second language learners (Sheorey & Mokhtari, 2001). This indicates that instruction should place greater emphasis on developing students' metacognitive knowledge and self-assessment capacity.

The results have significant pedagogical implications. Since retrieval strategies are essential to language comprehension, reading instruction should focus on activities that enhance inferencing, the use of context clues, and the activation of background knowledge. The poor scores on evaluative strategies further highlight the importance of formal, explicit self-evaluation practice to make learners more aware of their progress and to enable them to modulate their strategic decisions accordingly. Moreover, the observed stability and generalizability in strategy use imply that continued efforts should be made to advance refined, integrated-strategy instruction that aims to mediate both cognitive and metacognitive components, thereby supporting learners in becoming more flexible in applying strategies across tasks and contexts.

Overall, this study extends the current research literature by establishing a comprehensive SEM model that considers the dynamic relationship between cognitive and metacognitive strategies in EFL reading performance. The findings support the facilitating, though not central, role of strategies in L2 processing, provide further evidence for language knowledge as a dominant influence on this sort of task, and demonstrate that involvement in retrieval and monitoring is essential. The results extend our understanding of the strategic processing of L2 reading and offer pedagogical implications for strategy-based instruction in EFL settings.

VI. CONCLUSION

The present study used SEM to investigate the relationships between cognitive and metacognitive strategies and reading comprehension among Thai high school students. Results suggested that cognitive strategies were hierarchically arranged, with memory supporting retrieval, which, in turn, was the central mechanism underlying comprehension. This study indicates the importance of retrieval in constructing meaning from text. Metacognitive strategies, planning (planning in what you are going to say), monitoring (paying attention to one's speaking) and evaluation (assessing one's speaking), were also highly correlated; only the dimension of evaluation was identified as the least used by learners. Cognitive and metacognitive strategies significantly affected reading and tests through language knowledge, which mediated this effect in accordance with Bachman and Palmer's (1996) model. Taken together, these findings also underscore the importance of metacognitive regulation in the reading process and its influence on the strategic control of cognitive processes.

The results provide several practical and methodological insights. From a pedagogical point of view, they support the practice of explicit strategy-based instruction, which develops both cognitive and metacognitive processes. Instruction should focus on techniques that improve retrieval, such as activating background knowledge and inferring and using context clues. Still, learners should also be given tools to monitor their reading comprehension efforts more analytically.

The development of more conscious metacognitive awareness may also have implications for greater learner independence and greater effectiveness in dealing with texts. Methodologically, this research demonstrates the usefulness of SEM for more precisely capturing the interplay among reading variables and generates a testable model for further studies of strategic behaviour in reading.

Several limitations should be acknowledged. The sample was limited to Thai high school EFL students, so the results cannot be generalised directly to other situations, ages, or proficiency levels. Despite their accuracy, multiple-choice reading comprehension tests and Likert-scale questionnaires can fail to provide a complete analysis of reading comprehension or strategy use. Other test formats or task types may provide different insights. Additionally, the potential contribution of gender differences in strategy use and language development, widely documented across L2 research, could not be addressed with the available data, constraining a more complete description of learner variability.

A further direction for research is to investigate the impact of other assessment tasks (e.g., open-ended tasks, integrated reading and writing assessments, or think-aloud protocols) on language learners' strategy use and reading performance. Studies with a longitudinal, mixed-methods design might offer more insight into the development of these cognitive and metacognitive strategies. Future studies should examine gender and other language skills to better understand variation in strategy use across learner characteristics and proficiency domains. Lastly, intervention studies that systematically provide participants with training in evaluation and monitoring would help enhance strategic competence and promote autonomous learning in EFL settings.

ACKNOWLEDGEMENTS

This research project was financially supported by Mahasarakham University.

REFERENCES

- [1] Alderson, J. C. (2000). *Assessing Reading*. Cambridge University Press.
- [2] Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*, 75(4), 460-472. <https://doi.org/10.2307/329495>
- [3] Anderson, N. J., Bachman, L. F., Perkins, K., & Cohen, A. (1991). An exploratory study into the construct validity of a reading comprehension test: triangulation of data sources. *Language Testing*, 8(1), 41-66. <https://doi.org/10.1177/026553229100800104>
- [4] Anderson, N. J. (2008). *Practical English language teaching: Reading*. McGraw Hill.
- [5] Anderson, N. J. (2009). ACTIVE reading: The research base for a pedagogical approach in the reading classroom. In Z. H. Han & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 117-143). University of Michigan Press.
- [6] Anderson R. C., & Pearson P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *The handbook of reading research* (pp. 255-292). Longman.
- [7] Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language testing*, 17(1), 1-42. <https://doi.org/10.1177/026553220001700101>
- [8] Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- [9] Bachman, L. F., & Palmer, A. S. (2010). *Language Assessment in Practice: Developing language assessment and justifying their use in the real world*. Oxford University Press.
- [10] Bentler, P. M. (2006). *EQS structural equation program manual*. Multivariate Software.
- [11] Cai, Y., & Zhao, C. (2023). Metacognitive strategies and self-efficacy co-shape L2 achievement: A multilevel structural equation modeling approach. *System*, 117, 1-11. <https://doi.org/10.1016/j.system.2023.103099>
- [12] Cai, Y., King, R., & McInerney, D. M. (2022). The concurrent trajectories of utility value, metacognitive strategy use, and English achievement: A multivariate growth modeling analysis. *The Journal of Experimental Education*, 91(3), 472-493. <https://doi.org/10.1080/00220973.2022.2053496>
- [13] Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- [14] Cohen, A. D., & Upton, T. A. (2006). Strategies in responding to the new TOEFL reading tasks. *ETS Research Report Series*, 2006(1), i-162. <https://doi.org/10.1002/j.2333-8504.2006.tb02012.x>
- [15] Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- [16] Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- [17] Grabe, W., & Stoller, F. L. (2020). *Teaching and Researching Reading* (3rd ed.). Routledge.
- [18] Hair, J., Risher, J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2-24. <https://doi.org/10.1108/EBR-11-2018-0203>
- [19] Kitichaidateanan, P., & Sukying, A. (2025). Decoding success: Investigating the impact of trait and state strategies on reading test performance among Thai high school learners. *World Journal of English Language*, 15(1), 88-100. <https://doi.org/10.5430/wjel.v15n1p88>
- [20] Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.
- [21] Lin, L., Lam, J. W.-I., & Tse, S. K. (2019). Test takers' strategy use and L2 Chinese reading test performance in mainland China: A structural equation approach. *Studies in Educational Evaluation*, 60, 189-198. <https://doi.org/10.1016/j.stueduc.2019.01.002>
- [22] Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *Modern Language Journal*, 90(3), 320-337. <https://doi.org/10.1111/j.1540-4781.2006.00425.x>
- [23] McNamara, T. (1996). *Measuring second language performance*. Longman.

- [24] Ministry of Education of Thailand. (2008). *Revised indicators and learning standards for foreign languages in the Basic Education Core Curriculum B.E. 2551*. Retrieved January 9, 2025, from <http://www.act.ac.th/document/1741.pdf>
- [25] Mokhtari, K., & Reichard, C. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94(2), 249-259. <https://doi.org/10.1037/0022-0663.94.2.249>
- [26] Noipa, J., & Phusawisot, P. (2025). The effects of metacognitive reading strategy instruction on Thai EFL engineering students: Metacognitive strategy use and students' attitudes. *World Journal of English Language*, 15(2), 263-273. <https://doi.org/10.5430/wjel.v15n2p263>
- [27] O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.
- [28] Oxford, R. L. (2011). Strategies for learning a second or foreign language. *Language Teaching*, 44(2), 167-180. <https://doi.org/10.1017/S0261444810000492>
- [29] Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd ed.). Routledge.
- [30] Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language testing*, 20(1), 26-56. <https://doi.org/10.1191/0265532203lt243oa>
- [31] Phakiti, A. (2006). Modeling cognitive and metacognitive strategies and their relationships to EFL reading test performance. *Melbourne Papers in Language Testing*, 1, 53-96. <https://doi.org/10.3316/acipt.165763>
- [32] Phakiti, A. (2007). *Strategic competence and EFL reading test performance*. Peter Lang.
- [33] Phakiti, A. (2008). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language testing*, 25(2), 237-272. <https://doi.org/10.1177/0265532207086783>
- [34] Phakiti, A. (2016). Test takers' performance appraisals, appraisal calibration, and cognitive and metacognitive strategy use. *Language Assessment Quarterly*, 13(2), 75-108. <https://doi.org/10.1080/15434303.2016.1154555>
- [35] Purpura, J. E. (1997). An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning*, 47(2), 289-325. <https://doi.org/10.1111/0023-8333.91997009>
- [36] Purpura, J. E. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach*. Cambridge University Press.
- [37] Purpura, J. E. (2014). Cognition and language assessment. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1452-1476). Wiley-Blackwell.
- [38] Sari, M. I. (2016). Cognitive and metacognitive reading strategy use and reading comprehension performance of Indonesian EFL pre-service teachers. *Journal of Foreign Language Teaching and Learning*, 1(2), 46-61. <https://doi.org/10.18196/fil.1213>
- [39] Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons.
- [40] Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, 29, 431-449. [https://doi.org/10.1016/S0346-251X\(01\)00039-2](https://doi.org/10.1016/S0346-251X(01)00039-2)
- [41] Song, X. (2005). Language learner strategy use and English proficiency on the Michigan English Language Assessment Battery. *Spaan Fellow Working Papers in Second on Foreign Language Assessment*, 3, 1-26.
- [42] Song, X., & Cheng, L. (2006). Language learner strategy use and test performance of Chinese learners of English. *Language Assessment Quarterly*, 3(3), 243-266. https://doi.org/10.1207/s15434311laq0303_2
- [43] Sukying, A. (2021). Choices of language learning strategies and English proficiency of EFL university learners. *LEARN Journal: Language Education and Acquisition Research Network*, 14(2), 59-87. Retrieved March 23, 2025, from <https://so04.tci-thaijo.org/index.php/LEARN/article/view/253261>
- [44] Wenden, A. L. (1991). *Learner strategies for learner autonomy*. Prentice-Hall.
- [45] Zhang, L. (2018). *Metacognitive and cognitive strategy use in reading comprehension: A Structural Equation Modelling Approach*. Springer. <https://doi.org/10.1007/978-981-10-6325-1>
- [46] Zhang, L., Goh, C. M., & Kunnan, A. J. (2014). Analysis of test takers' metacognitive and cognitive strategy use and EFL reading test performance: A multi-sample SEM approach. *Language Assessment Quarterly*, 11(1), 76-102. <https://doi.org/10.1080/15434303.2013.853770>
- [47] Zhang, X., & Guo, L. (2020). Cognitive and Metacognitive Reading Strategies Training in EFL Reading. In *International Conference on Education, Economics and Information Management (ICEEIM 2019)* (pp. 110-114). <https://doi.org/10.2991/assehr.k.200401.028>

Panassanan Kitichaidateanan is a high school English teacher in Khon Kaen and a PhD student in the English Language Teaching Programme at Mahasarakham University, Thailand. Her academic interests include language testing, L2 reading, vocabulary acquisition, SLA, and teaching methodology, reflecting her commitment to advancing English language learning and teaching. Her email is muikitichai@gmail.com

Nithipong Yothachai works as a university lecturer in the English Education Program at the Faculty of Education and Human Development at Chaiyaphum Rajabhat University, Chaiyaphum, Thailand. His research interests include English language teaching, vocabulary acquisition, SLA, teaching methodology, academic writing, and linguistics. His email is nithiyo@gmail.com

Apisak Sukying (PhD) is an assistant professor and the Director of the PhD Program in English Language Teaching (ELT) at the Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand. He earned his PhD in TESOL from the University of Sydney, Australia. His interests include L2 vocabulary acquisition and development, L2 vocabulary assessment, SLA, learner strategies, and academic writing. His email is apisak.s@msu.ac.th

“I Am in Control”: Strategic Gender Performance and Agency in Joyce Carol Oates’s *Marya: A Life*

Lingling Xu*

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Malaysia

Manimangai Mani

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Malaysia

Hardev Kaur a/p Jujar Singh

Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Malaysia

Abstract—This article examines gender performativity as a strategic mode of agency in Joyce Carol Oates’s *Marya: A Life* (1986), focusing on the production of female scholarly identity within patriarchal academic institutions. Existing criticism has largely interpreted Marya’s development through psychological trauma, maternal reconciliation, or inner healing, tending to locate agency in psychic resolution rather than in the negotiation of institutional power. Drawing on Judith Butler’s theory of gender performativity, this study argues that Marya’s agency emerges through calculated gender performances rather than through the attainment of a coherent or authentic identity. Through close textual analysis, the article traces three interconnected stages in Marya’s trajectory. It first examines her adoption of masculinized behaviors as strategies for academic legitimacy, revealing the limits of masculinity when performed by a female body. It then analyzes her tactical switching between masculine assertion and feminine vulnerability in interpersonal and institutional encounters, showing how gender functions as a situational resource. Finally, it conceptualizes Marya’s later position as a feminist intellectual as a form of “meta-performance,” in which scholarly authority and feminist critique coexist within the same institutional framework. By foregrounding the performative construction of scholarly authority, this study reframes *Marya: A Life* as a critical exploration of how gendered agency is negotiated and rendered intelligible within patriarchal academic culture.

Index Terms—Joyce Carol Oates, *Marya: A Life*, gender performativity, feminist criticism

I. INTRODUCTION

As one of America’s most prolific contemporary writers, Joyce Carol Oates has long been recognized for her unflinching attention to violence, power, and gender in American life. Among her extensive works, *Marya: A Life* (1986) stands as a particularly compelling examination of female intellectual ambition and survival within patriarchal academic structures. Unlike Oates’s more Gothic or sensationalist works, *Marya: A Life* adopts a quieter bildungsroman form. The novel chronicles Marya’s struggles with poverty, her ambivalent relationship with femininity, her intense intellectual hunger, and her fraught romantic entanglements; most notably with her married academic mentor, Maximilian Fein. What distinguishes this novel is Oates’s nuanced portrayal of a female academic who refuses conventional feminine passivity while repeatedly confronting gendered limitations. Marya is neither victim nor heroine. Instead, she emerges as a complex figure who must constantly negotiate her identity to survive in spaces not designed for women like her. This ambiguity makes *Marya: A Life* an ideal text for examining how gender operates as a performed, strategic practice rather than a fixed identity.

II. LITERATURE REVIEW

Taken as a whole, existing scholarship reveals a sustained critical interest in Joyce Carol Oates’s engagement with American society and women’s experience, examined through psychological, historical, formal, and feminist lenses. These studies offer useful critical tools for reading individual works within her expansive oeuvre.

A. Past Studies on Joyce Carol Oates

Joyce Carol Oates has been the subject of sustained critical attention for several decades. Scholars have approached her work from a wide range of perspectives, responding in particular to what Johnson (1987) describes as her continuing engagement with “the phenomenon of contemporary America: its colliding social and economic forces, its philosophical contradictions, its wayward, often violent energies” (p. 8). Existing criticism on Oates may be broadly

* Corresponding Author. Email: gs61550@student.upm.edu.my

grouped into four interrelated approaches: psychological and spatial readings, historical and political criticism, postmodernist interpretations, and feminist analyses.

(a). *Psychological and Spatial Readings*

Early critical work on Oates tended to concentrate on questions of psychological disturbance and spatial organization in her fiction. Johnson (1987) observes that Oates portrays an America “in which individual lives are frequently subject to disorder, dislocation and extreme psychological turmoil” (p. 8). Approaching Oates from the tradition of tragedy, Grant (1978) argues that her protagonists construct a sense of identity through acts of self-striving (p. 159). Focusing on *The Rise of Life*, Tamilmani and Nagalakshmi (2021) read the characters’ fragmented psyches through the lens of postmodern textual space, linking this fragmentation to broader patterns of social decay (p. 242). Humann (2013) connects representations of domestic violence with adolescent agency, suggesting that Oates repeatedly presents trauma as a paradoxical site of empowerment rather than pure victimization (p. 84). Yuan (1993) examines female characters across lower-, middle-, and upper-middle-class positions, showing how patriarchal and class structures leave women economically marginalized and exposed to neurosis, violence, and even death.

(b). *Historical and Political Readings*

More recent criticism has turned to the historical and political dimensions of Oates’s writing. As Gates (1990) notes, “a future archeologist equipped only with her oeuvre could easily piece together the whole of postwar America” (p. 27). Oates herself has described her position as belonging to “the school of the writer as witness to history and society” (Germain, 1989, p. 177). Rather than reconstructing historical events directly, her fiction often attends to their aftermath, tracing the ways historical memory unsettles individual lives. Tromble (2015) shows how Oates incorporates moments of public crisis—such as the Chappaquiddick incident in *Black Water* (1992) and the Love Canal environmental disaster in *The Falls* (2004)—into narratives shaped by private trauma (p. 2). Recent studies also reflect growing interest in Oates’s political imagination, particularly her engagement with war. Olson (2022) exposes the disturbing logic of perpetual, profitable warfare and its repetitive cycles of trauma, while Bryla (2022) draws on the concept of liminality to read Oates’s depiction of Eastern Europe as a space that appears both familiar and alien, echoing Cold War power relations between the United States and the Soviet Union. In this vein, Yusupova et al. (2025) propose the concept of “Tragic America” in Oates’s fiction, contending that her use of violence and the portrayal of marginalized groups serve to deconstruct the illusion of the American Dream and expose deep-seated systemic inequalities (p. 1).

(c). *Postmodernist Interpretations*

Oates’s use of postmodern narrative strategies, such as metafiction, intertextuality, parody, fragmentation, and narrative decentering, has been a recurrent focus of critical discussion. Rather than adopting these techniques in a celebratory manner, critics generally note that Oates employs them to interrogate historical knowledge, cultural memory, and moral responsibility. McLennan (2020) argues that *The Accursed* simultaneously invokes and dismantles the conventions of historical romance, reading the novel as a cautionary reflection on America’s unresolved cultural anxieties entering the twenty-first century (p. 108). From a postsecular perspective, Ghosal (2020) interprets *The Accursed* as a reconfiguration of American history that fuses documented events with religious confession, thereby unsettling the boundary between faith and historical rationality (p. 14). Horváth’s (2022) study of *The Wonderland Quartet* shifts attention to aesthetic survival, suggesting that Oates’s fragmented narrative form responds to a world marked by the loss of transcendental meaning and echoes Nietzsche’s proclamation of the “death of God”. Miceli (2021) further emphasizes Oates’s formal experimentation by proposing that she adopts “a Fourth Way to tell the story” (p. 1) through the fictional memoir in *Black Water* (1992), *Blonde* (2000), and *My Sister, My Love* (2008). Taken together, these studies show that postmodern form in Oates’s fiction is inseparable from her ethical engagement with history and politics, rather than a purely aesthetic choice. Cologne-Brookes (2025) further observes that Oates’s postmodern “process-oriented” narratives often blur individual personality to challenge male-dominated discourses, suggesting that the meaning of her work remains a dynamic site of reader response rather than a fixed certainty.

(d). *Feminist Criticism*

Feminist interpretations of Oates’s fiction have undergone significant re-evaluation over the past four decades. Early scholarship, particularly prior to the 1990s, often hesitated to classify Oates as a feminist writer, largely because her early novels depict women as psychologically damaged, socially constrained, or subject to repeated violence (Gates, 1990, p. 28; Greene, 1991, p. 25; Zimmerman, 1986, p. 177). Goodman (1977) observes that female characters frequently attempt to reshape their destinies through marriage, sexuality, or motherhood, yet these strategies rarely result in meaningful liberation (p. 18). Daly (1996) offers a more developmental reading of Oates’s career, identifying shifts in authorial stance across decades: from “anxious authorship” in the 1960s, marked by violent father figures and threatened daughters (p. XVII), to “dialogic authorship” in the 1970s, when female characters begin negotiating agency (p. 70), and finally to “communal authorship” in the 1980s, as Oates constructs collective narrators and increasingly positions herself as a spokesperson for women’s experience (p. 125). By the 1980s, critical attitudes toward Oates’s feminism began to change. Showalter (1987) was among the first to recognize the feminist significance of Oates’s sustained attention to women’s lives, arguing that her fiction constitutes an important archive of female experience (pp.

137–142). Creighton (1992) similarly contends that Oates's later works foreground feminist concerns by examining how women are defined through male power and how they struggle to construct autonomous identities (p. 4). Parrott (1983) characterizes Oates as a progressive feminist who refuses to oppose women's liberation to motherhood, instead distinguishing between maternal identity and individual subjectivity in order to revise traditional gender relations (p. 2).

More focused feminist studies have addressed specific dimensions of gendered experience in Oates's fiction. Caliskan (2011) reads women's madness as a symptom of spiritual emptiness produced by a male-dominated capitalist society. Wesley's (1992) *Father-Daughter Incest as Social Transgression* examines *The Molesters*, them, and *Angel of Light*, arguing that incest functions not merely as personal trauma but as a transgressive disruption of patriarchal order. Creighton and Binette (2006) analyze mother-daughter relationships, suggesting that daughters' reinterpretations of their mothers' lives signal an emerging awareness of social inequality and gendered expectations. As Boileau and Tromble (2025) conclude, the literary representation of sexual violence in Oates's oeuvre is not a simple linear path from victimhood to empowerment, but a complex negotiation that mirrors the shifting political and cultural landscape of feminist thought over the last sixty years.

B. Past Studies on *Marya: A Life*

Critical studies of *Marya: A Life* have most often approached the novel through psychological frameworks. Within this body of work, the mother-daughter relationship is repeatedly identified as central to Marya's development. These studies offer valuable accounts of Marya's traumatic childhood and emotional formation. At the same time, they tend to interpret her agency primarily in terms of inner healing and psychic reconciliation, leaving her engagement with broader gendered structures less fully examined. Larson (1995), drawing on reader-response criticism and trauma theory as well as her personal experience, reads Marya's sadomasochistic tendencies as survival strategies shaped by domestic violence and childhood abuse. Larson (1995) argues that "adopting a masculine behavior alone will not enable her to find herself," and that reconciliation becomes possible only when Marya "acknowledge[s] and embrace[s] the hidden femininity and motherhood within her" (p. 79). Working within Freudian, Kleinian, and Lacanian frameworks, İleri (2009) likewise foregrounds the role of maternal love, concluding that reunion with the mother represents Marya's only route to inner peace. From this perspective, female empowerment emerges through a reaffirmed maternal bond that enables resistance to male-dominated society. Kealey (2015) follows a similar line of reasoning, suggesting that genuine self-understanding requires Marya to confront and accept her past, including her mother's presence within it. Approaching the novel from a different angle, Cologne-Brookes (2005) locates Marya's emotional reserve toward friends and colleagues in unresolved family trauma, arguing that her inability to achieve security and fulfillment stems from persistent inner conflict rather than from external constraints (p. 142).

Taken together, these psychologically oriented readings shed important light on Marya's interior life. Yet they also share a tendency to situate agency almost exclusively within the realm of psychic resolution including maternal reconciliation, acceptance of femininity, or the cessation of internal conflict. Such an emphasis risks downplaying the ways Marya actively negotiates male-dominated academic environments. By privileging interiority and familial repair, these studies often leave unaddressed how femininity and masculinity function as social codes that Marya learns to mobilize rather than inherent identities she must ultimately reconcile.

Nagalakshmi's (2021) study complicates this pattern by situating Marya within postmodernism and trauma theory, identifying her as "a postmodern survivor of desensitized familial and societal relationships in 1980s America" (p. 231). Nagalakshmi (2021) highlights two principal coping strategies: denial and emotional detachment from traumatic events, and the use of Catholicism and writing as forms of therapy. This approach productively brings social context into view and recognizes Marya's active efforts to survive. Even here, however, critical attention remains focused on coping and recovery rather than on the question of how Marya performs gender in order to obtain intellectual legitimacy or renegotiate power within academic and intimate relationships.

As a result, an important dimension of *Marya: A Life* remains insufficiently theorized. Although critics acknowledge the patriarchal constraints surrounding Marya, few explore gender performativity as a mode of agency in its own right. Little attention has been paid to Marya's tactical movement between masculine and feminine roles, including her calculated displays of vulnerability, her adoption of intellectual authority, or her manipulation of intimacy, as strategies shaped by institutional power. By reading *Marya: A Life* through Judith Butler's theory of performativity, the present study argues that Marya's gendered performances function not simply as mechanisms of survival, but as strategic interventions that expose the constructed nature of gender norms and open limited but consequential spaces for female agency within patriarchal academic culture.

III. METHODOLOGY

This study combines close textual analysis with feminist theory to examine how gender operates as a performative and strategic practice in Joyce Carol Oates's *Marya: A Life*. Grounded in Judith Butler's theory of gender performativity, the analysis treats gender not as an inner essence or stable identity, but as a series of acts shaped by social norms, institutional expectations, and relations of power.

Methodologically, it first considers Marya's early alignment with academic masculinity, intellectual discipline and emotional restraint as strategies to gain recognition within patriarchal institutions. The discussion then turns to moments

in which femininity operates not as an essential identity but as a tactical practice, with vulnerability and desire functioning as negotiated resources rather than signs of passivity. Finally, the study examines Marya's emergence as a feminist intellectual who works from within academic structures to interrogate their normative assumptions. Together, these readings frame gender performativity as a way to claim agency through which Marya navigates institutional constraint, contests authority, and produces critique without presuming a position outside power.

IV. RESULTS AND DISCUSSION

Building on Judith Butler's theory of gender performativity, the following discussion examines how Marya's agency takes shape through situated and strategic engagements with gendered norms. Instead of treating masculinity and femininity as stable positions, this analysis approaches gender as a repertoire of socially intelligible practices that Marya selectively mobilizes in response to institutional constraint. By tracing her movement from masculinized academic performance, through tactical gender switching, to a reflexive form of feminist scholarly authority, the discussion highlights how agency is negotiated within patriarchal academic structures.

A. Strategic Masculinization: Performing for Academic Legitimacy

Marya's engagement with masculinity first takes shape through her bodily difference, a physical presence that repeatedly places her at odds with conventional expectations of femininity. From early adolescence, Marya becomes acutely aware that her body and manner do not align with the feminine norms embodied by the other girls around her: "Marya was one of the tallest girls in the eighth grade: not precisely big-boned, not stocky, but solidly muscular, with long legs and narrow hips and alert nervous mannerisms" (Oates, 1986, p. 45). Butler (2011) argues that gender construction begins through performative language; when authorities say "This is a girl," it initiates what she calls "a setting of boundaries and an inculcation of norms" (p. xvii). To be named a "girl," in Butler's sense, is to be repeatedly trained into recognizable gender norms, learned not through explicit instruction but through everyday codes of appearance and comportment. However, Marya systematically refuses these normative performances. She remains completely indifferent to her appearance, unmoved even when a teacher comments that she looks "like a horse." Her physical presentation actively challenges stereotypical constructions of femininity:

Marya Knauer was too tall for her age, too long-waisted and angular; that her chest was flat; that her mouth reflected discontent when she believed herself in a neutral mood; that her cheekbones were too prominent, and her brows thick and dark and straight—indeed, they grew across the bridge of her nose and might almost have formed one single censorious brow if, in rage and self-loathing, she didn't pluck them every few weeks. (Oates, 1986, p. 75)

Her angular, unfeminine appearance does not simply mark difference; it gradually becomes the site through which she resists, and reworks, the meanings attached to femininity. Yet her "rage and self-loathing" during grooming reveals the internal tension of performing gender within regulatory structures. In high school, where "everything female depended upon being pretty or not-pretty" (Oates, 1986, p. 75), Marya's refusal to conform becomes more deliberate. She cuts her hair and wears short pants, which Marilyn identifies as symbolic rejection of "feminine identity" and "an attempt to evade victimization" (Oates, 1986, p. 128). She vows "not to be the kind of idiotic girl who worries constantly over her appearance," resisting becoming like her aunt Wilma, who "visibly warmed when paid a compliment" or "complained aloud about getting older".

Marya's masculine performance extends beyond appearance into behavioral and emotional realms. Her performances are always contingent upon situation and context. She appears invulnerable when encountering criticism from teachers, never crying in class like her "weaker classmates." Even at Father Shearing's hospital bedside, she restrains herself from assisting him with heavy books, as this would appear "too womanly, too solicitous" (Oates, 1986, p. 81):

Since the early days of her catechism class she had learned to suppress certain aspects of her personality in his presence: above all, any hint of weakness. She might be sinful in thought, word, and deed, but she was not going to be weak. (Oates, 1986, p. 77)

The sustained effort to suppress vulnerability suggests not emotional absence but a learned discipline, through which Marya fashions a posture of strength associated with masculine authority. Labeled "savage" by those around her, she embodies what Butler would describe as a failure to perform normative femininity, yet this failure is strategic. When her teacher Mr. Schwilk calls her a "savage beast" after she openly questions his authority, her defiance of the traditional submissive women results in punitive grading. Yet she confronts him directly, challenging his decision until he reluctantly raises her grade from B+ to A.

Marya's masculine performance extends beyond bodily demeanor into the intellectual sphere, where authority is tacitly aligned with masculinity. Her reading practices are telling: she immerses herself in an overwhelmingly male canon—"SCHOPENHAUER, DICKENS, MARX, Euripides. Oscar Wilde. Henry Adams. Sir Thomas More. Thomas Hobbes. And Shakespeare—of course" (Oates, 1986, p. 152). This list does not function merely as a record of intellectual curiosity; it marks Marya's early recognition that intellectual legitimacy is mediated through gendered traditions of authorship. That recognition surfaces explicitly when she reads Clifford Shearing's essay and pauses to wonder whether linguistic and intellectual proficiency might be "an entirely masculine skill, an art of combat by way of language, forever beyond her" (Oates, 1986, p. 95). The moment is less an expression of self-doubt than an

acknowledgement of the gendered economy in which knowledge circulates. Intellectual authority, as Marya comes to understand it, is not distributed neutrally but encoded through masculine norms that precede her entry into academic life.

Rather than withdrawing from this recognition, Marya responds by staging intellectual authority as a public performance. Her high school valedictory address—constructed through quotations from Thoreau, Emily Dickinson, Dostoyevsky, and Teilhard de Chardin—operates as a deliberate intervention into a gendered field of recognition. Standing on stage, she does not simply demonstrate learning; she claims visibility within a space where intellectual display remains implicitly masculine. The narrative underscores her awareness of the audience, describing her as relishing “the astonished gazes of men” as she deviates from expected norms of feminine modesty.

Emmett’s reaction crystallizes the stakes of this performance. His response is not admiration but injury: he feels that his superior male status had been offended, retreating into silence for days. What unsettles him is not Marya’s intelligence per se, but the ease with which she inhabits forms of intellectual authority tacitly reserved for men. Her performance exposes the fragility of masculine intellectual privilege precisely by demonstrating that its codes can be learned, cited, and enacted.

Marya’s immersion in the literary canon also sharpens her awareness of exclusion. She recognizes “the absence of women’s voices in literary tradition,” yet responds not with resignation but with what the narrative pointedly names as arrogance: “she told herself she would change all that” (Oates, 1986, p. 152). This ambition signals a refusal to treat intellectual space as a permanently male domain. At the same time, it reveals her belief that access might be secured through mastery of existing norms rather than their immediate transformation.

Read through Butler’s (1999) notion of cultural intelligibility, Marya’s intellectual performances can be understood as attempts to render herself legible as a legitimate knowing subject within male-dominated educational institutions. Importantly, intelligibility here does not denote mere visibility or competence, but recognition within an already established symbolic order—one that equates authority with masculinity even as it disavows doing so.

Yet Marya’s carefully staged masculine performances never result in full recognition. As Butler argues, those who fail to conform to recognizable gender norms are denied cultural intelligibility through exclusion and repudiation. Marya’s experience gives this abstraction a concrete social form. However fluently she performs intellectual masculinity, she remains marked as excessive and deviant—described as “horse-like,” “savage,” and fundamentally out of place within binary gender expectations. These descriptors do not simply insult Marya; they function as classificatory tools, reinscribing her difference as something unintelligible rather than oppositional. Her masculinity is not read as authority but as aberration, a misalignment between performance and body that cannot be resolved within existing norms.

In high school, even academic excellence fails to dislodge Marya from social determinism. She remains reducible to family identity—“Marya was a Knauer, after all” (Oates, 1986, p. 83)—and to gendered expectations that anticipate marriage rather than intellectual ambition. The contradiction at the heart of her position becomes increasingly apparent: performing femininity excludes her from intellectual legitimacy, yet performing masculinity renders her a failed woman, an incoherent figure who fits neither category without remainder.

This structural impossibility reveals the limits of singular performative strategies. Masculinity alone cannot secure Marya’s place within academic institutions because those institutions require not merely masculine performance but male embodiment. Her female body performing masculinity produces not recognition but cognitive dissonance, positioning her as the “savage” who cannot be assimilated into established classificatory schemes.

Marya’s persistent unintelligibility complicates any straightforward reading of gender performativity as a strategy of access. Rather than confirming the efficacy of masculine performance, her experience exposes its limits within institutions that conflate authority with embodied masculinity. Gender performance, in this context, does not function as a stable route to legitimacy but as a volatile and context-dependent negotiation.

It is precisely this failure that forces Marya toward more tactical deployments of gender. Unable to rely on masculinity as a singular solution, she must learn to shift between masculine and feminine codes in response to situational demands. Gender thus emerges not as a coherent identity to be perfected, but as a precarious resource to be managed—one that enables partial access even as it forecloses full recognition.

B. Subversive Repetition: Inverting Gender Hierarchies

The pen-stealing episode marks a moment in which Marya’s earlier reliance on masculine confrontation gives way to a more unstable and opportunistic negotiation of gendered codes. After Mr. Schwilk initially assigns her a punitive B+, Marya confronts him with directness and rational insistence, a mode of address aligned with masculine norms of authority. This strategy proves effective: Schwilk revises her grade to an A. Yet Marya does not rest with this apparent victory. Instead, she immediately shifts tactics, adopting a markedly different performative register by fabricating a narrative of maternal illness:

My mother is sick. She’s been sick all semester. I know I shouldn’t think about it much . . . I shouldn’t depress other people . . . but sometimes I can’t help it. She isn’t expected to live much longer, the cancer has metastasized to the brain. . . . I’m sorry if I offended you. (Oates, 1986, p. 158)

This performance risks reproducing familiar tropes of feminine vulnerability, yet Marya deploys them with acute awareness of their affective force. By presenting herself as a filial, suffering daughter, she mobilizes a culturally legible script designed to evoke Schwilk’s guilt and remorse. Crucially, this turn to vulnerability follows—rather than

replaces—her earlier masculine confrontation. The sequence reveals gender not as a fixed identity but as a repertoire of socially intelligible practices that can be activated situationally. Masculinity and femininity here operate less as stable positions than as provisional strategies, calibrated in response to shifting power relations.

Marya's subsequent theft of Schwilk's pen adds a further layer of symbolic resonance. Long associated with authorship, authority, and masculine intellectual lineage, the pen invites a symbolic reading that recalls Gilbert and Gubar's (1979) account of the phallic logic of writing. Marya's appropriation of the pen does not simply replicate masculine power; it occurs through the very performance of vulnerability that disarms Schwilk. Gender performance thus functions less as open resistance than as a covert and situational means of leverage, allowing Marya to extract symbolic authority without direct confrontation.

This episode complicates Butler's formulation of gender as both chosen and constrained. As Butler observes, "One chooses one's gender, but one does not choose it from a distance which signals an ontological juncture between the choosing agent and the chosen gender" (Butler, 1986, p. 40). Marya tactically selects which performance to deploy—masculine insistence or feminine fragility—yet both remain socially available scripts rather than free inventions. Her agency operates within, rather than outside of, patriarchal norms. At the same time, by juxtaposing and sequencing these performances, she denaturalizes gender, exposing it as what Butler elsewhere defines as "the repeated stylization of the body... within a highly regulatory frame" (Butler, 1999, p. 45) rather than an expression of innate being.

The instability of gendered authority becomes more explicit in Marya's sexual relationship with her mentor, Maximilian Fein. At its outset, the relationship reproduces conventional hierarchies: Marya occupies the position of assistant and lover, while Fein performs the role of the dominant male intellectual. Yet this structure gradually erodes as Fein's sexual failures puncture the coherence of his masculine authority. Oates describes these moments with unsettling physicality:

Sometimes he couldn't make love to her. Not in that way. He tried, he tried, half sobbing and clutching at her, pummeling her flesh, he tried and failed: she knew he despised her for his failure, she knew her breasts and belly and the insides of her thighs would carry the bruises for days. It doesn't matter, Marya consoled him, it doesn't matter. . . . But of course it did, to her aging lover it mattered greatly. (Oates, 1986, p. 218)

Read through a Lacanian framework, Fein's impotence signals a collapse of the phallus as a signifier of power, authority, and desire. His failure to perform masculine sexual dominance materializes the erosion of his symbolic position. Marya's response, however, is marked by performative complexity. Outwardly, she comforts him by saying "It doesn't matter" (p. 218). Internally, she recognizes that Fein's failure matters profoundly to him. This disjunction between performed reassurance and private awareness creates a space in which power quietly shifts.

Importantly, Marya's growing sense of control does not negate the violence enacted upon her body; rather, it emerges in uneasy proximity to it. Her agency develops not in the absence of harm but alongside it, complicating any straightforward narrative of empowerment. As Butler suggests, such disjunctions open space "for new, subversive performances of gender and sexuality," where repetition introduces difference rather than stability.

Marya always encourages herself by saying "MARYA WAS ALWAYS thinking, These days, these weeks, I am in control of the situation. I am in control, in control, of the situation" (Oates, 1986, p. 219), which marks a moment of conscious appropriation. This is not a rejection of the heterosexual script but a reworking of it through what Butler terms subversive repetition. Marya continues to perform the outward signs of feminine submission while internally occupying a position of control. In doing so, she unsettles the presumed alignment between masculinity and dominance, femininity and passivity, revealing these roles as contingent rather than natural.

Marya's negotiation of gendered power finds its most overtly physical expression in the cycling scene with her colleague Gregory. Historically, the bicycle has functioned as a symbol of female mobility and emancipation—a "freedom machine" that challenged norms of feminine dependence (Wånggren, 2017, p. 124). Yet it has also been a site of anxiety, with women cyclists often framed as gender deviants or sexual transgressors (Joans, 2001, p. 142). As Marya and Gregory await tenure decisions, their shared ride initially reproduces internalized hierarchies. Marya measures her body against Gregory's, admiring his "thin but strong legs" while doubting her own physical capacity. The male body functions as an implicit standard, reinforcing what Butler identifies as the performative maintenance of bodily hierarchies. Gender here is sustained not through explicit prohibition but through habitual comparison and self-surveillance.

This dynamic begins to fracture as the ride progresses. Marya's physical exertion—marked by anxiety, exhilaration, and "a muscular sort of rage"—condenses her professional frustration and her resentment toward male authority. When she reaches a downhill stretch and allows herself to accelerate, the scene shifts register:

With a thrill of satisfaction she knew she was speeding faster than she'd ever sped before; and accelerating every second... Now you can all go to hell, she thought. (Oates, 1986, p. 264)

The act of speeding becomes a symbolic rejection of the male figures who once disciplined and intimidated her. When Gregory suggests stopping to rest, Marya refuses, asserting her endurance against his caution. In this reversal, Oates destabilizes heteronormative binaries that associate masculinity with risk-taking and femininity with restraint. Marya's recklessness and Gregory's conservatism expose these traits as performative conventions rather than natural dispositions.

Rather than functioning as a phallus per se, the bicycle enables Marya to access forms of mobility, risk, and bodily assertion historically aligned with masculine power. Through physical performance, she claims authority over movement and danger without mediation by male approval. The cycling scene thus extends Marya's tactical use of gender into the realm of embodied action, demonstrating that even physical capacity is neither biologically fixed nor symbolically neutral.

C. Meta-Performance: Constructing Feminist Scholarly Identity

Marya's later intellectual practices cannot be separated from the bodily and relational strategies through which she first learned to negotiate authority. Early in her academic career, she actively suppresses non-academic writing, fearing that it might expose a "sensitive nature" culturally coded as feminine and therefore incompatible with scholarly authority. Her hesitation reflects an internalized understanding of legitimate academic voice as detached, impersonal, and implicitly masculine.

While serving as Fein's assistant, Marya publishes a non-academic essay in the prestigious magazine *The Meridian*. This publication marks a tentative attempt to insert a personal voice into public intellectual discourse. However, when she shares the piece with Fein, "hoping for his approval," his response is marked by disdain rather than encouragement. Deeply affected by his indifference, she declines the editor's invitation for further contributions. At this stage, Marya remains largely compliant, rarely challenging Fein's authority or asserting her own intellectual priorities.

At the same time, this period of apparent subordination quietly equips her with institutional knowledge unavailable through formal training. Acting as Fein's assistant grants her access to academic networks, critical vocabularies, and professional practices typically monopolized by male scholars. Rather than functioning solely as a position of dependency, the assistantship becomes a site where Marya accumulates intellectual capital while outwardly conforming to hierarchical expectations. Compliance, in this context, operates less as submission than as a provisional tactic.

When Marya leaves her teaching position and relocates to New York, she trades institutional security for professional uncertainty. Her choice—cultural criticism and journalism over a stable academic post—performs the risk-taking and ambition that academic institutions often discourage in women. Like her earlier bicycle ride, this geographic relocation registers as refusal, signaling her unwillingness to remain within the subordinate institutional roles typically available to female academics.

Marya's trajectory, however, is not one of permanent exit. She neither fully assimilates into academic norms nor rejects them outright. Instead, she leaves when institutional structures constrain her agency too tightly and later returns when she can occupy a position that affords greater discursive authority. This movement away and back underscores a strategic negotiation with institutions rather than a binary choice between complicity and resistance.

By the time Marya assumes the role of feminist scholar lecturing on contemporary feminist theory, her performative strategies have undergone a qualitative transformation. Whereas she previously deployed masculine or feminine performances to navigate specific institutional contexts—performing masculinity to gain academic legitimacy and tactically switching between gender scripts to manipulate power relations—her scholarly position turns performativity itself into an object of analysis. Her position thus operates as a reflexive mode of performance in which enactment and critique coincide. She occupies an institutional position whose authority depends on norms she explicitly interrogates; rather than merely navigating gendered expectations, she teaches their construction. Crucially, meta-performance designates not a higher or purer form of agency, but a specific configuration in which scholarly authority and feminist critique are mutually constitutive. Marya performs academic authority—a role historically coded as masculine—while arguing that such authority is produced through repeated institutional practices rather than grounded in inherent intellectual legitimacy.

The analytical force of this configuration lies in its specificity. Because Marya's authority as a lecturer depends on credentials, institutional recognition, and professional legitimacy—all conditions shaped by male-dominated institutional structures—her critique emerges from within the very framework it interrogates. As Kim (2007) suggests, identities may sustain conventions while simultaneously generating subversive effects. Marya participates in academic life—publishing, lecturing, holding institutional roles—while exposing the gendered assumptions that organize those practices.

This reflexivity becomes especially visible when the content of Marya's teaching is read alongside her biography. Her lectures on feminist theory presumably address how women are positioned by patriarchal discourse, how subordination is reproduced through normative repetition, and how female intellectuals navigate male-dominated institutions. These are not abstract problems for Marya; they are the conditions through which her own authority has been negotiated. She lectures on the very processes she has lived through and continues to enact. Teaching feminist theory thus becomes inseparable from performing the role of "feminist scholar," a role attained only through earlier tactical accommodations to gender norms. When she discusses Butler's theories of gender performativity, she simultaneously occupies a position of scholarly authority while drawing attention to the conditions that made that position intelligible in the first place.

In this sense, Marya's position resonates with Butler's (1997) description of subjectivity as "a site of . . . ambivalence in which the subject emerges both as the effect of a prior power and as the condition of possibility for a radically conditioned form of agency" (pp. 14–15). Marya's identity as a feminist intellectual is shaped by institutional power: her position exists because patriarchal academia has authorized such roles, often as circumscribed spaces for women's

voices. Simultaneously, she uses that position to contest institutional power by making its operations visible. Performance and critique remain entangled; there is no external vantage point from which she can speak. Meta-performance names this paradox without resolving it, insisting that situated, embedded critique remains both constrained and consequential.

V. CONCLUSION

Marya's development traces a progression in how gender performativity functions as a resource for agency within patriarchal institutions. Initially, she adopts masculine-coded behaviors to gain intellectual legitimacy, only to encounter persistent unintelligibility. She then learns to switch tactically between masculine and feminine performances to negotiate power in interpersonal and institutional contexts. Finally, she occupies a scholarly position that enables her to examine gender performativity from within academic authority itself.

This study contributes to feminist readings of Oates by shifting attention from gendered behavior alone to the construction of scholarly identity as a performative process. Marya's significance lies not simply in her resistance to masculine dominance, but in how her authority as a feminist intellectual is produced through—and constrained by—the same institutional norms she critiques. Her agency does not transcend structure; it operates through calculated engagement with it.

Rather than offering a narrative of either liberation or defeat, Marya's trajectory illustrates how feminist agency in academia takes shape as ongoing negotiation. Her meta-performative position as a feminist scholar does not dismantle patriarchal structures from the outside, but inhabits them critically, exploiting their instabilities. This progression—from masculine performance through tactical switching to meta-performance—reveals a crucial insight: feminist intervention often proves effective not through refusing institutional roles, but through inhabiting them in ways that expose their performative construction. In doing so, Oates's novel presents a nuanced account of how gendered authority is both reproduced and contested within academic institutions, framing performativity not merely as a theory of gender, but as a constitutive condition of intellectual authority in patriarchal academic culture.

REFERENCES

- [1] Boileau, N. P., & Tromble, T. L. (2025). Exploring specters of feminism in the work of Joyce Carol Oates. *Bearing Witness: Joyce Carol Oates Studies*, 6(1), Article 1.
- [2] Bryla, M. (2022). A feast for the Cold-War imagination: Liminal Eastern Europe in the writings of John Updike, Joyce Carol Oates and Philip Roth. *European Review*, 30(4), 478–489.
- [3] Butler, J. (1986). Sex and gender in Simone de Beauvoir's *Second Sex*. *Yale French Studies*, 72, 35–49.
- [4] Butler, J. (1997). *The psychic life of power: Theories in subjection*. Stanford University Press.
- [5] Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity* (2nd ed.). Routledge.
- [6] Butler, J. (2011). *Bodies that matter: On the discursive limits of "sex"* (2nd ed.). Routledge.
- [7] Caliskan, D. Ü. (2011, February 17). An existentialist study: Suicide as a consequence of madness in Joyce Carol Oates's fiction. In *International Conference on Management, Economics and Social Sciences (ICMESS2011)* (pp. 621–624).
- [8] Cologne-Brookes, G. (2005). *Dark eyes on America: The novels of Joyce Carol Oates*. Louisiana State University Press.
- [9] Cologne-Brookes, G. (2025). Joyce Carol Oates and feminism: Facts found and foundered. *Bearing Witness: Joyce Carol Oates Studies*, 6(1), Article 2.
- [10] Creighton, J. V. (1992). *Joyce Carol Oates: Novels of the middle years*. Twayne Publishers.
- [11] Creighton, J. V., & Binette, K. (2006). "What does it mean to be a woman?": The daughter's story in Oates's novels. *Studies in the Novel*, 38(4), 440–456.
- [12] Daly, B. (1996). *Lavish self-divisions: The novels of Joyce Carol Oates*. University Press of Mississippi.
- [13] Gates, H. L., Jr. (1990, July 2). Murder she wrote: Review of *Because it is bitter, and because it is my heart*. *The Nation*, 27–29.
- [14] Germain, D. (1989). Interview with Joyce Carol Oates. In L. Milazzo (Ed.), *Conversations with Joyce Carol Oates* (pp. 177–185). University Press of Mississippi.
- [15] Ghosal, N. (2020). Gothicizing American history: Religion, race, and politics in Joyce Carol Oates' *The Accursed*. *The IUP Journal of English Studies*, 15(2), 14–24.
- [16] Gilbert, S. M., & Gubar, S. (1979). *The madwoman in the attic: The woman writer and the nineteenth-century literary imagination*. Yale University Press.
- [17] Goodman, C. (1977). Women and madness in the fiction of Joyce Carol Oates. *Women and Literature*, 5(2), 17–28.
- [18] Grant, M. K. (1978). *The tragic vision of Joyce Carol Oates*. Duke University Press.
- [19] Greene, G. (1991). *Changing the story: Feminist fiction and the tradition*. Indiana University Press.
- [20] Horváth, D. O. (2022). *Nietzsche and Joyce Carol Oates: Nietzschean themes in The Wonderland Quartet*. Academica Press.
- [21] Humann, H. D. (2013). Domestic violence, child agency, and the adolescent perspective in Joyce Carol Oates's *Freaky Green Eyes*. *Children's Literature Association Quarterly*, 38(1), 84–93.
- [22] İleri, F. (2009). *The familial trauma and search for parents in Joyce Carol Oates's Marya: A Life* [Doctoral dissertation]. Dokuz Eylül University.
- [23] Joans, B. (2001). *Bike Lust: Harleys, women, and American society*. University of Wisconsin Press.
- [24] Johnson, G. (1987). *Understanding Joyce Carol Oates*. University of South Carolina Press.
- [25] Kealey, J. T. M. (2015). "I know you!": The implications of knowing in Joyce Carol Oates's *Marya: A Life*. *Bearing Witness: Joyce Carol Oates Studies*, 2(5).

- [26] Kim, D. K. (2007). The agency that difference makes: Judith Butler's theory of performativity and the irony of the political. In *Melancholic freedom: Agency and the spirit of politics* (pp. 83–106). Oxford University Press.
- [27] Larson, A. L. (1995). *Marya: A Life, by Joyce Carol Oates: Masochistic behavior as a recovery strategy of the abuse survivor* [Master's thesis]. Iowa State University.
- [28] McLennan, R. (2020). "We cannot create": The limits of history in Joyce Carol Oates's *The Accursed*. In R. Maxey (Ed.), *21st Century US Historical Fiction: Contemporary Responses to the Past* (pp. 95–110). Palgrave.
- [29] Miceli, B. (2021). A "fourth" way to tell the story: Fact and fiction in three novels by Joyce Carol Oates. Peter Lang.
- [30] Nagalakshmi, R. (2021). Narratives of survival: invisible oppression and retaliation in Joyce Carol Oates's *Marya: A Life*. *British and American Studies*, 27, 231–239.
- [31] Oates, J. C. (1986). *Marya: A life*. E. P. Dutton.
- [32] Olson, D. (2022). Joyce Carol Oates' *Carthage* (2014): death, maidens, & revenant witnesses. In *Gothic War on Terror: Killing, Haunting, and PTSD in American Film, Fiction, Comics, and Video Games*, Cham (pp. 109–134). Springer International Publishing.
- [33] Parrott, M. L. M. (1983). *Subversive conformity: Feminism and motherhood in Joyce Carol Oates* [Doctoral dissertation]. University of Maryland, College Park.
- [34] Showalter, E. (1987). Joyce Carol Oates: A portrait. In H. Bloom (Ed.), *Modern critical views: Joyce Carol Oates* (pp. 137–142). Chelsea House.
- [35] Showalter, E. (1999). *A literature of their own: British women novelists from Brontë to Lessing*. Princeton University Press.
- [36] Tamilmani, K. T., & Nagalakshmi, R. (2021). Dismemberment of Kathleen's psyche in Joyce Carol Oates's *The rise of life on Earth*. *English Studies at NBU*, 7(2), 227–244.
- [37] Tromble, T. L. (2015). Fiction in fact and fact in fiction in the writing of Joyce Carol Oates. *Bearing Witness: Joyce Carol Oates Studies*, 2(1), Article 2.
- [38] Wänggren, L. (2017). *Gender, technology and the New Woman*. Edinburgh University Press.
- [39] Wesley, M. C. (1992). Father-daughter incest as social transgression: A feminist reading of Joyce Carol Oates. *Women's Studies*, 21(3), 251–263.
- [40] Yuan, W. C. (1993). *Images of women in selected works of Joyce Carol Oates* [Doctoral dissertation]. State University of New York at Buffalo.
- [41] Yusupova, H. O., & Bekova, G. I. (2025). Tragic America in the novels of Joyce Carol Oates. *Journal of Applied Science and Social Science*, 15(02), 24–28.
- [42] Zimmerman, B. (1986). Feminist fiction and the postmodern challenge. In L. McCaffery (Ed.), *Post-modern fiction: A bibliographical guide* (pp. 175–188). Greenwood Press.



Lingling Xu was born in Huaian, China, in August 1982. She received the M.A. degree in English Language and Literature from Yangzhou University, Yangzhou, China. She is currently pursuing the Ph.D. degree in English literature with the Faculty of Modern Languages and Communication, Universiti Putra Malaysia (UPM), Serdang, Malaysia.

She is currently a Lecturer with Huaiyin Institute of Technology, Huai'an, China. She has authored or co-authored several papers in academic journals, including *When Objects Speak: Thing Narrative, Trauma, and Identity in Joyce Carol Oates's Carthage* (2025). Her research interests include American literature and feminist literature. Email: gs61550@student.upm.edu.my

Manimangai Mani is a Senior Lecturer at the Department of English, Faculty of Modern Languages and Communication, Universiti Putra Malaysia; her research interests include Post-colonial Studies, African and Caribbean Literature. E-mail: manimangai@upm.edu.my

Hardev Kaur is an Associate Professor at the Department of English, Faculty of Modern Languages and Communication, Universiti Putra Malaysia; her research interests include English literary studies, African and Caribbean literature studies, and Trauma Studies. E-mail: hardevkaur@upm.edu.my

Fear-Driven Dialogue: A Linguistic Study of *Death Knocks* and *Death*

Mohammed A. Abou Adel*

Department of Arabic Language and Literature, College of Humanities and Social Sciences, UAE University, Al Ain, UAE

Ayman I. Elhalafawy

English Language and Literature at Faculty of Arts, Kafrelsheikh University, Kafrelsheikh, Egypt

Samir H. Khalifa

Medical Sciences and Preparatory Year Department, North Private College of Nursing, Arar, Saudi Arabia

Ahmad M. Al Mahamed

Department of Languages, College of Arts, Education and Social Sciences, Abu Dhabi University, Abu Dhabi, UAE

Fatma T. El-Zaghal

English Language and Linguistics at Faculty of Arts, Tanta University, Tanta, Egypt

Abstract—This study explores how fear as a mood state affects conversational patterns and shapes communicative behavior. It offers a novel perspective on how fear influences both verbal and nonverbal interactions. Drawing on dialogues from Woody Allen's *Death Knocks* and *Death*, the analysis employs Conversation Analysis by Jack Sidnell (2010) and the multimodal extensions proposed by Lorenza Mondada (2019). The study addresses a notable gap by applying both Conversation Analysis (CA) and Multimodal Conversation Analysis (MMCA) to Woody Allen's modern drama, where gestures, posture, and movement are central to meaning-making. The study also moves beyond single-discipline approaches by adopting an interdisciplinary framework that brings together applied linguistics, theatre studies, and emotional psychology. This allows for a deeper understanding of how theatrical discourse is shaped under the pressure of existential emotion and offers a genuine contribution at the intersection of these knowledge domains. The findings indicate that fear disrupts logical conversation, resulting in brief, tense exchanges characterized by emotional rigidity and cognitive imbalance. The study contributes to improving dialogue management in real-life interactions and dramatic scripts, particularly in contexts of psychological tension. Future research may examine how other emotions—such as joy, anger, or anxiety—shape conversation to deepen our understanding of the relationship between emotion and communication.

Index Terms—multimodal conversation analysis, Fear of Death, Woody Allen, *Death Knocks*, *Death*

I. INTRODUCTION

Communication is a vital aspect of our daily lives, and the conversational patterns we follow play an important role in determining the success of our interactions. By being aware of verbal and nonverbal cues and respecting cultural differences, one can enhance communication skills. Communication patterns may vary depending on the context, culture, and the purpose of communication (Shahbar, 2024). Many factors affect communication styles, such as the psychological state of the speaker, the motivation behind the dialogue and mental pressures. According to Uddin (2023), communication barriers can hinder our ability to communicate effectively, leading to misunderstandings and missed opportunities. Identifying these barriers becomes the first step in overcoming them. Communication barriers, generally, may include language barriers, cultural barriers, and psychological barriers. If the speaker does not speak in the listener's own language, it may lead to confusion and misunderstanding (Yuldasheva, 2024).

According to Mazeland (2006), conversation analysis (CA) has made a growing contribution to a number of established domains in applied linguistics. The study of conversation analysis is incorporated into more traditional linguistic fields like discourse analysis, pragmatics, or (interactional) sociolinguistics. One can comprehend meaning, power relationships, and social structures that are ingrained in spoken language through conversation analysis (Antaki, 2008; Adel, 2019). Methods of conversation analysis vary, and for many linguists such as Wooffitt (2008), Sidnell (2010), and Clift (2016), conversational aspects include turn taking, repair, and adjacency pairs in addition to openings and closings. In order to accomplish specific objectives, the two involved parties must alternate roles during the communication process. The most common elements in CA include analyzing data itself, being aware of the context of conversation, categorizing data in order to specify the best method for analyzing conversation, and visualization, which is latent in predictive analysis.

This study belongs to the field of applied linguistics, particularly within the area of discourse and conversation analysis, adopting a multimodal approach to investigate interactional meaning-making. Although numerous studies have examined fear from psychological perspectives, they have overlooked the dynamic relationship between emotion and language. Moreover, most studies have treated dramatic dialogue as a purely textual phenomenon, disregarding the multimodal aspects that shape communication on stage. This gap underscores the need for an analytical framework that integrates verbal and nonverbal modes to capture the full complexity of fear in interaction. Accordingly, the central research question asks: How does the fear of death shape, disrupt, or transform communication patterns in Woody Allen's *Death Knocks* and *Death*?

The methodology examines real-time interaction between characters, attending to fine details of conversational organization, including pauses, gestures, intonation, and body movement. This study differs from previous work by combining verbal and nonverbal analysis of theatrical dialogue and by focusing specifically on fear as a psychological state rather than on general character or script analysis. It also demonstrates Woody Allen's ability to challenge traditional views of death by presenting it in a relatable, often humorous, manner.

II. LITERATURE REVIEW

Recent scholarship has significantly broadened the scope of Multimodal Conversation Analysis (MMCA), offering valuable tools for studying how speech, gesture, gaze, and movement work together in interaction. Van de Mierop (2020) demonstrates how multimodal cues influence participation and turn-taking, revealing that conversational authority often stems from embodied signals rather than words alone. Doehler et al. (2021) similarly show how speakers use subtle actions—such as hesitations, shifts in gaze, or changes in posture—to manage delicate responses and negotiate interpersonal alignment. Such developments have been supported by new technological tools like ConAn (Penzkofer & Müller, 2021), which strengthens the precision of multimodal transcription. Further contributions from Lilja (2022) and Pellet-Rostaing et al. (2023) highlight how individuals coordinate linguistic, visual, and affective resources in constructing meaning and engagement, while Tomasine (2024) illustrates the role of multimodality in managing multiple activities during assessment tasks. Alqahtani (2024) adds another dimension by examining how power relations are enacted through language and embodied conduct. More recently, MMCA research has moved into new contexts: Jang et al. (2025) explore immersive multimodal systems that integrate gesture and gaze into chatbot communication; Hassan et al. (2025) analyze how spatial positioning structures safety dialogues; Zhou (2025) investigates identity construction in multimodal English conversations; and Chang et al. (2025) outline a methodological model for capturing the structural organization of multimodal interaction. Collectively, these studies showcase MMCA's growing sophistication and reinforce its relevance for analyzing complex, embodied communication such as theatrical performance.

In parallel, literary and linguistic scholarship underscores the importance of integrating text-based and performance-based approaches when analyzing dramatic dialogue, particularly in emotionally charged contexts. Alhourani et al. (2025a) demonstrate how linking linguistic analysis with theatrical expression uncovers deeper cultural and emotional layers, as shown in their study of identity in Malik's *Unveiled*. Moreover, Alhourani et al.'s (2025b) related work on Parks' *In the Blood* emphasizes how trauma shapes communicative behavior, highlighting the strong connection between psychological states and conversational patterns—an insight that aligns closely with the present study's focus on fear. Elhalafaway et al. (2025) further argue that gesture, voice quality, and stage movement are central to meaning-making, reinforcing the need for a multimodal approach when examining dramatic texts. Together, these contributions affirm the value of combining Conversation Analysis with MMCA to understand how emotions—especially fear—structure dialogue, influence interactional rhythm, and shape characters' communicative behavior in Woody Allen's *Death Knocks* and *Death*.

III. METHODOLOGY

This study adopts a qualitative approach within the framework of Conversation Analysis (CA) as presented by Sidnell (2010) and expanded by Mondada (2019) through a multimodal perspective. It examines how fear as a psychological state shapes verbal and non-verbal interaction in dramatic dialogue. CA provides the tools to analyze how fear disrupts, delays, or restructures conversational flow, while the qualitative design allows in-depth exploration of turn-taking, meaning-making, and multimodal resources. The study focuses on two one-act plays, *Death Knocks* and *Death*, which center on the theme of death, offering a rich context to analyze fear in both speech and embodied behavior.

Sidnell's (2010) CA framework guides the analysis of conversation openings and closings, adjacency pairs, turn-taking, TCUs, sequential organization, preference structures, and repairs. Mondada's (2019) multimodal approach incorporates multimodality, multisensoriality, temporal organization, embodied orientation, sequentiality of actions, and the material environment, highlighting how bodily actions co-construct meaning. Dialogues are analyzed according to CA conventions, with attention to stage directions and nonverbal cues. Scenes dominated by anxiety and fear are segmented and coded, examining pauses, hesitations, overlaps, interruptions, sequence shifts, repairs, gestures, gaze, sudden movements, freezing, and spatial changes. Stage directions in the published texts provide the primary source for non-verbal features. The analysis proceeds through familiarization, scene segmentation, and coding, allowing the researcher to interpret how fear affects conversational coherence, rhythm, and structure, and how embodied actions reinforce or counter verbal

expressions. By situating dialogue within its multimodal performance context, this methodology moves beyond purely linguistic analysis, capturing the embodied enactment of mood states such as fear.

IV. DISCUSSION AND FINDINGS

In this study, the researcher attempts to shed light on the conversational patterns in some selected plays written by Woody Allen. She endeavors to elucidate the significance of context in shaping the conversational pattern. The context here is not only limited to the linguistic one, but also refers to the message behind the events of each play and the psychological health that prompts each character to arrange the conversational pattern in different ways. The researcher has chosen Woody Allen's plays because he is the first playwright to embody the fear of death in most of his plays and show its effect on the way characters organize the dialogue. Also, Woody Allen's plays are characterized by the smoothness of the dialogue, being rich in all types of dialogue such as pauses, overlaps, backchannels, adjacency pairs, preference, turn-taking, and repair.

Conversational Patterns in *Death Knocks* (1968)

Death Knocks is a play by Woody Allen. It takes place only between two characters: Nat, the tragic hero, and Death. The play is considered a series of unconventional conversations, especially in its opening, because the prevailing psychological atmosphere in the play is Thanatophobia. In analyzing that play, the researcher examines the mechanisms of social interaction, concentrating on the micro-level aspects of talk. The dialogue between the two characters in this play is designed in a disjointed manner. It is full of pauses, reflecting the state of tension and dread associated with the fear of death. The repetition of certain phrases and the overlapping of voices reveal an inner anxiety, making fear seem present as a hidden third party in the dialogue.

Dialogue 1

Nat: Who are you?

Death: Death

Nat: Who?

Death: Death. Can I sit down?

I'm shaking like a leaf.

Nat: Who are you?

Death: Death. You got a glass of water?

Nat: Death? What do you mean, Death?

Death: Is it Halloween?

Nat: No

Death: Then I'm Death.

Now can I get a glass of water or a Fresca? (Allen, 1968, p. 1)



Figure 1. A Scene From *Death Knocks*
(Joshi, 2020, 00:00:57)

At the beginning of the play, Death horribly appears to Nat. As a result, Nat feels afraid, and his speech becomes disorganized. According to Afzali (2013), not all human beings believe in death, even when faced with it, considering themselves more powerful than it. Out of Nat's fear, he uses rising intonation and single words, repeating his question to Death, "Who are you?" three times without taking pauses. With his quick tone, Nat uses unrelated short answers; the length of the conversation is not fixed between them. He uses stress on the interrogative word "who" and talks before Death reaches possible completion. According to Levinson (2012), concerning "mapping actions," where each turn in conversation maps out actions, Death appears for a particular task which is ending the life of Nat. Although Death should be more dominant out of his position and status as an angel of death, the order of speaking isn't decided in advance.

Nat dominates the wheel of conversation from the moment Death appears. Fear and unexpectedness lead to a deliberate

violation of the natural conversational rules. According to Al-Mamoory and Al-Ghizzy (2023), when displaying communication in pairs of utterances, openings and closings are crucial components. Conversational openings and closings employ adjacency pairs. Examples are "Hello," "How are you?" "See you later," and "Bye." Although Death visits Nat for the first time, the exchange does not follow the conventional opening sequence; their dialogue begins without any greeting or pre-sequence. Greetings typically establish mutual orientation between participants at the beginning of any dialogue.

Turn-taking is the fundamental mechanism that regulates conversational flow, ensuring that speakers alternate smoothly and cooperatively to maintain mutual understanding (Meyer, 2023). Nat's way of turn-taking reflects his attempt to take control of the situation through his rapid and short questions. The conversation is also not dominated by asymmetrical adjacency pairs out of fear; each question does not receive the expected response. Repair is latent in formulation and clarification caused by disbelief. Also, for preference and sequential organization, the structure of the conversation reveals a deliberate deviation from the expected social responses. The sequential organization is not linear; it is characterized by disruption followed by reorganization. Death asks Nat to get a glass of water or a Fresca and to sit down, but Nat ignores his request. One's expectation of adjacent pairs in the form of request-response has been left stranded. Nat ignores Death's request. Also, when Death asks Nat questions, Nat's response does not display full conviction but rather a reluctant alignment with Death's claim, shaped by fear and uncertainty. Nat's second response, saying "no" reflects a correct adjacent pair when Death asks, "Is it Halloween?" For Death, he starts to repair the overlap, clarify his identity, and give answers to Nat's questions. Death says, "Then I'm Death".

Fear is also reflected upon Nat's body language. He resorts to nonverbal resources such as freezing and stiffness, wide open eyes, nervous gestures, forward lean, and vocal hesitation to express his astonishment. His tense, uneasy gaze and the way he opens his hands subtly convey fear, tension, and a sense of disbelief. The sensory environment is embodied through the calm voice, the dim lighting, and the trembling tone. The emotional atmosphere is perceptible through multiple sensory channels. Multiple temporalities coexist in this scene; Death's speech and movements unfold timelessly since he is associated with eternity. Nat's reactions are hurried, and this contrast creates the dramatic tension between life and death. Nat dominates the spatial composition, having control over the scene.

For materiality, the embodied presence of the two characters creates meaning. Death's dark robe materializes power and Nat's physical withdrawal embodies vulnerability. According to Mondada (2019), in social interaction, materiality refers to the way tools, objects, and spaces are active resources, not passive backgrounds since participants engage with material resources. Each material element has an organizational force in the sense that it facilitates possibilities for interaction. In moments of fear, moving away from an object or escaping from loud noises reorganizes spatial/material relations. Hence, materiality translates fear into physical actions and supports interaction.

Dialogue 2

Nat: Now, wait a minute

Death: I'm sorry. I can't help you.

Nat: You look a little like me.

Give me some time.

Death: I can't. What do you want me to say?

Nat: One more day. You play chess?

Death: No, I don't.

Nat: You play gin rummy?

Death: Do I play gun rummy?

Nat: You're good, huh?

Death: Very good.

Nat: I'll tell you what I'll do (Allen, 1968, p. 4)



Figure 2. A Scene From *Death Knocks*
(Joshi, 2020, 00:04:37)

In dialogue 2 topic shift is motivated by fear. Both Nat and Death take turns, but for the second time, Nat breaks the natural flow of the conversation and changes the pace of the dialogue by talking about another topic to calm down. Nat tells Death that he looks like him, and this is considered irrelevant regarding the main context of the dialogue. Nat constantly interrupts Death and reiterates his desire not to go with him. Nat's state of confusion urges him not to give a chance for Death to respond. He keeps his verbal turn longer than usual. He prolongs the dialogue as if it could give him a better chance to live. Both Nat and Death use adjacency pairs of request and refusal. Death says, "I'm sorry," and "I can't." Rejecting requests is a dispreferred response in the CA preference strategy. Nat then asks Death to play with him, but Death doesn't refuse that time. These adjacency pairs are unbalanced as Nat asks many questions and Death gives short answers. This behaviour reflects an asymmetry in power relations between them.

Nat begins to take pauses before speaking to mark interruptions and overlaps. When Nat decides to exploit his intelligence and forget his fear, the pattern of conversation changes since his psychological state has become stable. For preference, Death's refusal to Nat's requests from the beginning reflects unfavorable responses. Nat's new request to play games to save his life reflects his escalating tension and fear. Nat uses repair to defuse the situation or redirect the conversation. Upon refusing Nat's request, Nat repairs the critical situation by turning it into a friendly game. Death himself uses repair when he says, "Do I play gun rummy?" It is an example of repair initiation by misunderstanding. In talking about playing, Nat asks Death whether he is good at playing or not, saying, "You are good, huh?" "Huh" as a backchannel signal denotes asking. For sequential organization, this dialogue progresses from rejection and pleading to negotiation, then playing and bargaining. This is considered a natural sequence under the pressure of fear.

Body language and nonverbal cues complete the verbal dialogue, enriching its meaning. According to statistics given by the University of California, 7% of human communication is limited to words only, while 55% depends on body language, movements, and gestures, and 38% on the tone of voice and the pauses used. Body language is the language that is most responsible for effective human communication (Almotery et al., 2024). Nat starts to stand up to confront Death as an attempt to save his life. He tries to regain control by extending his hand over Death's face as a sign of challenge. Nat tries to create a safe distance in confronting Death through words and movements. His fear begins to transform from physical paralysis to defensive activity. Nat's tone of voice changes during defence and persuasion from a trembling tone to a calm, persuasive one.

Dialogue 3

Nat: So I'll see you tomorrow.

Death: What do you mean you'll see me tomorrow?

Nat: I won the extra day. Leave me alone.

Death: You were serious?

Nat: We made a deal.

Death: Yeah, but

Nat: Don't "but" me. I won twenty-four hours.

Come back tomorrow.

Death: I didn't know we were actually playing for time.

Nat: That's too bad for you. You should pay attention. (Allen, 1968, p. 7)



Figure 3. A Scene From *Death Knocks*
(Joshi, 2020, 00:09:34)

This dialogue reflects how conversational patterns and social behavior differ with the difference of the psychological states of the involved participants. When Nat starts to play, he exchanges his serious tone for a playful one. By forgetting fear, he stops repeating words and questions, replies directly, and explains. Turn-taking in this dialogue takes place quickly; there are no periods of silence, and the situation this time is going in Nat's favor. Nat achieves success over Death, defeats him, takes the initiative, and interrupts Death. Nat's fear drives him to speak quickly before Death's refusal to their deal regarding the game. The process of turn taking between Nat and Death is presented in turn construction units

(TCU) where they use different sentence structures, including sentences, clauses, phrases, and lexical items. They anticipate completion to TCU, targeting this point as a suitable place for starting a new talk. For sequence organization, the dialogue progresses from explanation to insistence and justification. Each sentence is stronger than the previous one. Adjacency pairs used in this dialogue show the conflict between realistic perception (Death) and defensive illusion (Nat). Nat's fear makes him give a response that is logically inconsistent but psychologically consistent with his desire to survive. Death starts to use derogatory questions and back channel signals like "yeah." He feels stressed out of defeat, repeating Nat's words in a denouncing manner.

Since Nat feels at ease upon forgetting the idea of death, his speech starts to be organized. Nat starts to use complete sentences, give orders and dispreferred responses in responding to Death to show his refusal to Death's authority. He says, "Come back tomorrow," "Leave me alone," and "See you tomorrow." By contrast, Death starts to feel interrupted upon his failure, starting to ask more derogatory questions such as "you were serious?" and "from what account?" He uses backchannel signals such as "yeah, but..." Moreover, like Nat at the beginning of the conversation, he repeats the interrogative word "what?" out of surprise. His speech starts to be random and disorderly. Death just tries to repair the misunderstanding, but Nat refuses this repair, saying, "You should pay attention".

For paralinguistic features, Nat's tone in imperative sentences reflects triumph and agitation at the same time. His body language appears strongly defensive. Nat's expanded fingers and open postures become an attempt to show dominance and impose his perspective. Spreading fingers here does not suggest self-confidence but rather an attempt to impose authority. For Death, he appears weaker than Nat in this scene, and his joined fingers suggest indifference and indicate a restrained, evaluative stance. Death speaks calmly, surrendering to a reality that he finds unacceptable. His body appears closed, suggesting a loss of control as if he is in a defensive position. Hence, all movements appear synchronized with the chosen words, and the interaction displays an embodied asymmetry.

Dialogue between characters in this play does not adhere to the main dialogue strategies; it lacks a natural conversational opening in the opposite of what happens between characters when they meet for the first time. "Topic shift" as a conversational strategy is employed when Nat exploits the game between him and Death to escape the predicament in which he is placed. Also, certain strategies are used, such as preferred and unpreferred responses, requests for clarification, exchange of roles, and correction of what is said. The closing stage of the conversation doesn't include the use of descending intonation or an atmosphere of politeness and comfort between the two parties. Since fear is the main theme in this play, Nat's fear makes him be involved in more activities, including speech and bodily perceptions. In talking to Death, objects that Nat is occupied with are sensory entities with tangible properties, such as his glasses, his newspaper, and the playing cards. He interacts with these objects to hold on to reality and to forget fear. He sometimes withdraws from verbal interaction to concentrate on other senses.

Conversational Patterns in *Death* (1975)

Death deals with a thorny topic that Woody Allen previously raised in 1968, which is escaping from death. The play revolves around Kleinman, who is awakened by a group of people who throw him into a futile pursuit of the serial murderer, and he finds himself forced to face this fate. During the chase, Kleinman meets several characters who participate in a plan that is unknown to him, and each of them gives their opinion on the matter of the chase. While the group is busy searching for the murderer, Kleinman finds himself face-to-face with the murderer. Fear of death is the main theme in the two selected plays. There is a similarity between the openings of the two plays: *Death* and *Death Knocks*, and thus the conversational patterns in both plays seem similar.

Dialogue 1

Al: We need you. Get dressed.

Kleinman: What?

Sam: Let's go.

Kleinman: What is this?

Al: Come on, move.

Kleinman: Move where? It is the middle of the night.

John: The killer's been spotted. They saw him entering the park.

Kleinman: What killer?

Victor: This is no time for babbling.

Kleinman: Who is babbling? What killer? I'm in a deep sleep? (Allen, 1975, p. 48)



Figure 4. A Scene From Woody Allen's One Act Plays *Death & God*
(Slater, 2019, 00:07:20)

Unlike *Death Knocks*, in *Death*, conversations take place between various characters who seem afraid of confronting death, and this is reflected in their way of speaking. The play starts with the appearance of Kleinman, with his rising intonation and a drop in pitch. Upon hearing knocks on the door, Kleinman starts to employ short signals of surprise in the form of a question to denote disbelief, confusion, and surprise, such as "huh?" It is used as an inquiry for receiving an affirmative reply. The strangers who come to Kleinman at midnight give orders from the moment they see him without using greetings. They say, "Hey," as an interjection to call attention. Kleinman repeats his question word "what," so an atmosphere of turbulence prevails. He also uses exclamation marks and more interrogative questions, such as "which maniac?" and "which strangler (Allen, 1975, p. 49)?" Kleinman repeats the declarative sentence "I'm coming" (p. 46) and replies to their orders with condemnation or ignorance, and this is a dispreferred response concerning CA. All of them are talking quickly, and Kleinman keeps interrupting them.

In dialogue one, since the strangers fail to choose an appropriate setting, the dialogue seems pointless. It doesn't include the two primary phases of any ordinary dialogue: the opening stage and the closing stage. It lacks greetings, falling intonation, formulaic expressions, discourse markers, and pre-closings. Turn-taking doesn't proceed naturally. The strangers take turns and give orders to Kleinman, giving him no chance to understand. They don't adhere to the main rules governing turn-taking. Kleinman's role becomes restricted to expressing astonishment, hesitations, and asking questions. However, the strangers don't focus on any arranged set of repair techniques. Despite referring to the killer as John, Kleinman asks, "What killer?" So, the trouble source lies in not understanding what happens.

The stage of "repair initiation" starts when Kleinman asks many questions for clarification. Starting to provide information about the killer is the "repair outcome." Kleinman's insistence on asking many questions for clarification is considered "a second repair initiation." Kleinman then interjects and breaks in, so the interruption seems competitive. He only begins to show politeness once he starts to understand. He says, "You can relax (p. 51)." However, when they begin to suspect him, Kleinman starts to use requests, silence, more declarative sentences, rising intonation, and long pauses to show innocence. As a result, they begin to depend on backchannel signals such as "uh-huh" to show understanding. One of them begins to show anger, asking him not to ask questions again. He also uses interjections such as "well" before his requests to introduce a remark.

For action sequencing and nonverbal cues, the sequence of actions is disrupted; the dialogue doesn't follow a request-response pattern. There is a series of overlapping sequences. This reflects the lack of shared understanding. Concerning preference organization, all responses seem dispreferred and face-threatening for Kleinman, and Kleinman's responses also seem dispreferred for the strangers. Kleinman shows intense fear, shock, and a feeling of helplessness by putting his hands on his head as a way to protect his body mentally and physically. He shows astonishment towards an unbearable reality. All involved characters in this dialogue fail to foster understanding, build relations, or communicate well out of fear.

Dialogue 2

Kleinman: Doctor!

Doctor: I'm dying.

Kleinman: I'll get a doctor.

Doctor: I am a doctor.

Kleinman: Yes, but you are a dying doctor.

Doctor: It's too late he caught me... ugh...

There was no place to run.

Kleinman: Help! Help! Somebody come quickly!

Did you get a good look at him?

Doctor: No, just suddenly, a stab in the back.

I'm dying, Kleinman. (Allen, 1975, p. 82)



Figure 5. A Scene From Woody Allen's One Act Plays *Death & God*
(Slater, 2019, 00:33:38)

The dialogue begins unnaturally between Kleinman and the doctor, as is common in plays where the emotional context influences the conversational pattern. Kleinman expresses his exclamation upon seeing the doctor dying. There is a natural and sequential system of turn-taking in this dialogue. Both the doctor and Kleinman take turns without any interruption from the other. Due to the high tension and fear, Kleinman expands the role and holds the floor for a longer period, expressing the internal state of anxiety. Kleinman starts to direct two questions to the doctor to change the pace of the dialogue, which is considered a set of declarative sentences from the beginning. The adjacency pairs illustrate the dynamic interactions between the statement and response, the statement and correction, and the question and answer.

In this dialogue, there are two types of repair mechanisms. In saying, "I'm a doctor," the doctor performs self-initiated self-repair, and in saying "yes, but you are a dying doctor," Kleinman performs other-initiated other-repair. "The trouble source" is the point where both Kleinman and the doctor signal the problem. The moment when they start to repair each other's utterances is called "repair initiation." The moment when they appear cooperative, accepting each other's repair attempts, is called "the outcome." Regarding repair mechanisms, the speaker performs a self-repair in a self-initiated manner. The recipient performs a repair in self-initiated other-repair. Other-initiated self-repair occurs when the recipient signals a problem in the speaker's speech, and, in return, the speaker clarifies the utterance. In other-initiated other repair, the recipient signals the problem and corrects it (Mills & Redeker, 2023). Out of repairing each other's responses, most responses are dispreferred ones in the context of CA. However, this doesn't suggest any disagreement between the two parties; they just provide clarification. Regarding sequence organization, the sequence begins with shocking information and ends with a dramatic escalation. Each utterance builds on the previous one and contributes to the escalation of the situation until the tragic end of the doctor's death.

Nonverbal cues, including prosodic and paralinguistic ones, help in embodying the atmosphere of horror in this scene. The vocalized particle "ugh" used by the doctor, pauses, overlaps, and the falling intonation reveal both fear and annoyance. It is part of turn- turn-constructive units presented by Sidnell (2010). It fills the gap during thinking or denotes moments of realization before completing speech. Pauses taken by the doctor don't mean a sequence break, but they are used as a sign of a repair initiation coming. Multiple sensory cues such as sight, touch, gestures, and facial expressions play a vital role in conveying the psychological atmosphere of the scene. The faint, trembling voice, the dying breath, and the dying body are all material resources that express the suffocation of time and space. Therefore, materiality contributes to the creation of fear in this scene. Kleinman's touch on the doctor's body and the glances exchanged between them constitute both physical and tactile dimensions of multimodal interaction. This dialogue is considered a series of intertwined temporal events. The scene reveals the time of the external event, the time of the internal feeling of fear, and the time of the inevitable end. The enclosed space, darkness, and the lack of an escape route are all material elements that create the psychological pressure.

Dialogue 3

Maniac: Kleinman?

Kleinman: Who are you?

Maniac: The homicidal killer.

Can I sit down? I am exhausted.

Kleinman: You are---the killer?

Maniac: Sure. Don't get excited.

I am armed.

Kleinman: You---You are going to kill me?

Maniac: Of course. That's my specialty.

Kleinman: You---You are crazy.

Maniac: Sure I am crazy. (Allen, 1975, p. 101)



Figure 6. A Scene From Woody Allen's One Act Plays *Death & God*
(Slater, 2019, 00:49:51)

Fear reaches its climax in the third dialogue as Kleinman meets the killer for the first time, and this naturally reflects on the pattern of conversation. Maniac starts talking to Kleinman without using greetings; he uses elliptical questions, and Kleinman, as a reaction, uses information-seeking questions. Out of Kleinman's fear, he adopts long pauses, delays, repetition, and falling intonation. His speech in this dialogue is restricted to rhetorical interrogative questions; he doesn't use any declarative sentences. Turns are organized in sequence, and there are no overlaps. Maniac dominates turn-taking, while Kleinman only asks questions to seek information. Kleinman's turns are responsive. This pattern shows asymmetry in turn distribution. However, in reaction to Maniac's request, "Can I sit down?" Kleinman ignores responding to it; he disrupts the turn-taking system. This is considered a dispreferred action displacement or sequence disruption.

Adjacency pairs are core units of conversation, linking an initial action with a predictable response to maintain coherence and manage interaction (Nonomura & Mori, 2025). All adjacency pairs used in the form of questions and answers show deviation from preferred responses since Maniac always provides confrontational answers, not socially unfamiliar ones. Hofstetter (2020) explains that each action recalls another relevant one as in "questioning" and "answering." Giving a reaction becomes a sign of understanding. For preference organization and repair mechanisms, Maniac uses dispreferred responses to create a dominant identity. This is called "dispreferred turn-format," as Sidnell (2010) puts it. Also, in reaction to Kleinman's questions, Maniac does not make any self-repair or clarification; he only creates deliberate interactional trouble. Regarding sequence organization, this dialogue progresses through predictable sequences. The sequence develops gradually towards a dramatic climax. Regarding action formation, Kleinman adopts topic shift in responding to the request for permission presented by Maniac. Vocal tones of the two characters support the ascending sequence of the conversation.

Like all previous dialogues, non-verbal communication plays a decisive role in conveying messages. Kleinman puts his hand on his cheek, covering his eyes as a non-verbal cue expressing intense fear. This frequently occurs in Woody Allen's plays. It is a trial for protecting oneself and hiding feelings of fear. Maniac puts his hand on Kleinman's shoulder, and this gesture suggests threat and physical control. Since looking directly at the camera creates a direct connection with the audience, both Maniac and Kleinman look at the audience with a scary look to convey psychological fear. Multisensoriality serves as a means of conveying feelings of control and fear. For multiple temporalities, different actions continue, such as threats, while others are instant, such as sudden replies. Mondada (2019) asserts that each interactive moment is connected to its time, overlapping with others to form the complete meaning. In this dialogue, lighting, close distance between characters, body itself, and the place as a space taken for interaction are considered interactive resources. Dim lighting, for example, denotes horror.

Dialogue 4

John: Kleinman... what was he like?

Kleinman: Like me.

John: What do you mean, like you?

Kleinman: Oooh... You gonna listen to Jensen
or you gonna listen to me?

John: All right, don't get angry...

Kleinman: All right, then, don't talk like a jerk...
He looked like me...

John: Bring him some water.

Kleinman: What do I need water for?

John: Are you afraid to die?

Kleinman: It's not that I'm afraid to die,
I just don't want to be there when it happens.

Oh... oh... ugggmmff. (Dies) (Allen, 1975, p. 106)



Figure 7. A Scene From Woody Allen's One Act Plays *Death & God*
(Slater, 2019, 00:53:16)

In this dialogue, there are many characters involved in the scene, but two characters are only talking to each other: John, the police officer, and Kleinman. Without greetings, the officer asks many questions and gives commands. Kleinman's responses are short and defensive since he is dying. The change-of-state token "Oooh" is used by Kleinman to express resentment. It functions as a pause to turn the dialogue in his favor and to disrupt the natural flow of dialogue. There is no natural turn-taking sequence; there is a shift from a question-answer sequence to a question-question sequence in replying, "You gonna listen to Jensen....?" Then, the pattern of adjacency pairs changes to a request and a protest question. Kleinman says, "What do I need water for?" After that, a natural turn-taking sequence of a question and answer appears. Some contextualization cues, such as "Oooh" used by Kleinman and "All right" used by John, reflect John's cooperation and Kleinman's resentment.

Since repair mechanisms are related to authority and agitation, different kinds of repair appear in this dialogue. In asking, "What do you mean, like you?" John makes other-initiated self-repair. Since Kleinman replies wrathfully, the situation returns to repair failure. Kleinman, then, makes self-initiated self-repair when he clarifies the matter, saying, "It is not that....?" Regarding preference organization, most responses are dispreferred, especially Kleinman's ones in protesting against John's questions. Also, instead of thanking John for ordering some water for him, he provides a contradictory response. In addition, in responding to John's question, "Are you afraid to die?" Kleinman gives a detailed response instead of giving answers with either "yes" or "no," and this reflects his hesitation and imbalance. John, by contrast, gives a preferred response in responding to Kleinman's question, saying, "All right, don't get angry." Fear creates progressively escalating stress in the sequence of this dialogue. According to Pillet-Shore (2017), preference organization organizes a variety of activities, including agreements and disagreements, acceptances and declines. Preferred turns are face-affirming, whereas dispreferred ones are face-threatening. While dispreferred second pair-parts are typically delayed or mitigated, preferred second pair-parts are produced promptly and without delay.

In interaction, Mondada (2019) calls for focusing on the senses, as people experience the world through bodily and sensory perceptions. Objects and materials are therefore not merely tools to be used but sensory entities with tangible properties that participants orient to within interaction. The tactile sensation in this dialogue plays a vital role in conveying meaning, and facial expressions are not clear. All the investigators lean towards Kleinman with pity. There is no physical distance between all characters. John's tone is low, reflecting his anxiety. Likewise, Kleinman's tone of voice sounds low, reflecting his surrender, fear, and inner tension. The chronological sequence creates a gradual escalation of fear. The investigators' body language reflects an attempt to intervene and protect Kleinman, while Kleinman's hunched body on the stage with his hands on his knees reflects pain and surrender.

V. CONCLUSION

Through study and analysis, it becomes clear that in moments of fear, conversational patterns undergo noticeable shifts, as speakers tend to modify their linguistic choices, turn-taking behaviors, and response timing in order to manage heightened emotional pressure. Such changes often manifest through shorter utterances, delayed responses, increased interruptions, and a greater reliance on nonverbal cues, all of which reflect the speaker's attempt to regain control over the interaction. The concept of death functions as a powerful psychological and linguistic force capable of reshaping the structure, tone, and dynamics of human dialogue. Fear of death not only alters lexical choices and rhetorical strategies but also drives speakers towards heightened emotional expression, reduced cognitive elaboration, and increased reliance on affiliative, defensive, or ritualistic forms of language.

Moreover, the study proves that fear of death functions as a powerful force, restructuring human dialogue across several interconnected dimensions. Existential fear impacts the embodied and spatial dimensions of communication. Gestures, posture, vocal qualities, and spatial positioning reveal heightened anxiety and serve as mechanisms for emotional regulation or symbolic stability. Dialogue under fear is compressed, metaphorically intensified, and marked by shifts in

turn-taking, politeness, and coherence. These patterns demonstrate that communication under existential threat is both a cognitive negotiation and a survival strategy. By integrating linguistic and psychological perspectives, the study highlights how fear transforms dialogue into a multisensory, temporal, spatial, and embodied experience. The findings underscore the value of interdisciplinary approaches to understanding extreme emotional states and their implications for clinical communication, crisis interactions, and interpersonal discourse.

REFERENCES

- [1] Adel, M. A. (2019). The semiotics of paratexts in Haifa Betar's novel *Bleeding Dreams*: A study in modern literary criticism. *An-Najah University Journal for Research – B (Humanities)*, 33(6), 963–972. <https://doi.org/10.35552/0247-033-006-003>
- [2] Alhourani et al. (2025a). Decolonizing the Stage: Exploring Postcolonial Narratives and Identity in Rohina Malik's *Unveiled*. *World Journal of English Language*. <https://doi.org/10.5430/wjel.v16n1p95>
- [3] Alhourani, M. et al. (2025b). Beyond the Mother Knot: Trauma and Intersectionality in Suzan-Lori Parks' *In the Blood*. *Theory & Practice in Language Studies (TPLS)*, 15(8). <https://doi.org/10.17507/tpls.1508.05>
- [4] Afzali, K. (2013). Character revelation and dialogue interpretation through Politeness theory and Conversation analysis in dramatic discourse: The case of Woody Allen's *Death Knocks*. *Journal of Language and Literature*, 2(6), 43-54.
- [5] Allen, W. (1968). *Death knocks*. The New Yorker.
- [6] Allen, W. (1975). *Without Feathers*. New York: Random House.
- [7] Al-Mamoory, S. et al. (2023). A pragmatic study of turn-taking and adjacency pairs in online conversations. *International Journal of English Language Studies*, 5(2), 84–93. <https://doi.org/10.32996/ijels.2023.5.2.8>
- [8] Almotery, A. et al. (2024). Bridging Communication Gaps: A Study on Effective Patient Communication Among Respiratory Therapy Students and Interns. *Cureus*, 16(5). <https://doi.org/10.7759/cureus.60484>
- [9] Alqahtani, D. (2024). Conversation Analysis: The Case of the Power of Language. *Egyptian Journal of Linguistics and Translation*, 12(1), 174-193. <https://doi.org/10.21608/ejlt.2023.238051.1041>
- [10] Antaki, C. (2008). Discourse analysis and conversation analysis. *The SAGE handbook of social research methods*, 431-446. <https://doi.org/10.4135/9781446212165.n25>
- [11] Chang, K. K. et al. (2025). *Multimodal conversation structure understanding*. arXiv. Accessed 23rd November, 2025, from <https://arxiv.org/abs/2505.17536>
- [12] Clift, R. (2016). *Conversation analysis*. Cambridge University Press. <https://doi.org/10.1017/9781139022767>
- [13] Doehler, S. et al. (2021). Multimodal Assemblies for Prefacing a Dispreferred Response: A cross-linguistic Analysis. *Frontiers in Psychology*, 12, 689275. <https://doi.org/10.3389/fpsyg.2021.689275>
- [14] Elhalafaway, A. et al. (2025). From Text to Performance: A Lehmannian Study of Postdramatic Strategies in Churchill's *Escape Alone* and Stephens' *Light Falls*. *Journal of Language Teaching and Research*, 6, 1824-1832. <https://doi.org/10.17507/jltr.1606.04>
- [15] Hassan, S. et al. (2025). Coherence-driven multimodal safety dialogue with active learning for Embodied Agents. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025)* (pp. 950-958). International Foundation for Autonomous Agents and Multiagent Systems. <https://doi.org/10.5555/3709347.3743614>
- [16] Hofstetter, E. (2020). Sequence Organization: Understanding What Drives Talk. In A. De Fina & A. Georgakopoulou (Eds.), *The Cambridge Handbook of Discourse Studies* (pp. 121–142). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108348195.007>
- [17] Jang, J. et al. (2025). *Enabling Chatbots with Eyes and Ears: An Immersive Multimodal Conversation System for Dynamic Interactions*. arXiv, pre-print arXiv:2506.0042. <https://doi.org/10.18653/v1/2025.acl-long.1519>
- [18] Joshi, A. (2020, January 30). *Death Knocks (a Play by Woody Allen) -The Drama Club, BITS Goa*. [Video]. YouTube. <https://www.youtube.com/watch?v=aTFgacD3k0Q>. Accessed 27th December, 2025.
- [19] Levinson, S. C. (2012). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103–130). Malden, MA: Wiley-Blackwell. <https://doi.org/10.1002/9781118325001.ch6>
- [20] Lilja, N. (2022). Multimodal conversation analysis as a method for studying second language use and learning in naturally occurring interaction. In *Gesture and multimodality in second language acquisition* (pp. 204-228). Routledge. <https://doi.org/10.4324/9781003100683-9>
- [21] Mazeland, H. (2006). Conversation analysis. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 3, pp. 153–162). Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00314-X>
- [22] Meyer, A. S. (2023). Timing in conversation. *Journal of Cognition*, 6(1), 1-17. <https://doi.org/10.5334/joc.268>
- [23] Mills, G. J., & Redeker, G. (2023). Self-Repair Increases Referential Coordination. *Cognitive Science*, 47(5). <https://doi.org/10.1111/cogs.13329>
- [24] Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment, materiality ,multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62. <https://doi.org/10.1016/j.pragma.2019.01.016>
- [25] Nonomura, R., & Mori, H. (2025). Who speaks next? Multi-party AI discussion leveraging the systematics of turn-taking in Murder Mystery games. *Frontiers in Artificial Intelligence*, 8, 1582287. <https://doi.org/10.3389/frai.2025.1582287>
- [26] Pellet-Rostaing, A. et al. (2023). A multimodal engagement model integrating conversational, visual, and affective signals. *Frontiers in Computer Science*, 5, 112-130.
- [27] Pillet-Shore, D. (2017). Preference organization. In J. Nussbaum (Ed.), *The Oxford research encyclopedia of communication*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.132>
- [28] Penzkofer, A. et al. (2021). ConAn: A usable tool for multimodal conversation analysis. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21)* (pp. 341–351). ACM. <https://doi.org/10.1145/3462244.3479886>
- [29] Shahbar, H. (2024). Developing intercultural competence: The role of communication in the adaptation of Saudi women in Canada. *Open Journal of Social Sciences*, 12, 595–613. <https://doi.org/10.4236/jss.2024.1212039>
- [30] Sidnell, J. (2010). *Conversation analysis: An introduction*. Oxford: Wiley-Blackwell.

- [31] Sidnell, J. (2020). *Conversation Analysis*. Retrieved November 16, 2025, from <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003031758-14/conversation-analysis-eva-maria-martika-jack-sidnell>.
- [32] Slater, A. (2019, November 2). *Woody Allen's one-act plays DEATH & GOD* [Video]. YouTube. <https://www.youtube.com/watch?v=N-TGtq9Z7Tc>. Accessed 27th of Dec 2025.
- [33] Tomasine, J. S. (2024). Multiactivity during formal, formative reading assessment: A multimodal approach. *Frontiers in Communication*, 9, 134-150. <https://doi.org/10.3389/fcomm.2024.134150>
- [34] Uddin, M. A. (2023). Communication barriers in work environment: Understanding impact and challenges. *International Journal of Research and Innovation in Social Science*, 7(11), 1201–1206. <https://doi.org/10.47772/IJRISS.2023.71118>
- [35] Van De Microop, D. et al. (2020). Investigating the interplay between formal and informal leaders in a shared leadership configuration: A multimodal conversation analytical study. *Human Relations*, 73(4), 490-515. <https://emcawiki.net/VanDeMicroop-et-al2020>. Accessed 27th December, 2025.
- [36] Wooffitt, R. (2008). *Conversation analysis* (2nd ed.). Polity Press.
- [37] Yuldasheva, D. K. (2024). Communication barriers in intercultural communication and ways to overcome them. *European ScienceMethodical Journal*, 2(8), 67–73. <https://www.researchgate.net/publication/392135971>. Accessed 27th December, 2025.
- [38] Zhou, Y. (2025). An analysis of English multimodal conversation based on identity construction. *CLEC 2025 Proceedings*, 3, 78-85. https://doi.org/10.2991/978-2-38476-430-3_20



Mohammad A. Abu Adel is a scholar specializing in modern Arabic literature. He holds a PhD with honors in Modern Literature (2014), an MA (2009), and a BA (2005) from Damascus University. With 14 years of teaching experience, he has served as an associate professor at Al Yamamah University (2019–2024) and currently teaches at UAE University, UAE. Dr. Abu Adel has authored over thirty publications in globally indexed journals. He is an active member of academic committees, including Course Equivalency and Accreditation. Email: mohamad.abouadel@uaeu.ac.ae & <https://orcid.org/0000-0001-7987-5464>



Ayman Elhalafawy is a professor of English Language and Literature and head of Department of English Language at Faculty of Arts, Kafrelsheikh University, Egypt. Prof. Elhalafawy's key areas of research include Minority Literature, Globalization and Literature, posthuman literature, and American Literature. Email: ayman.elhalafawi@art.kfs.edu.eg & <https://orcid.org/0009-0000-9115-685X>



Samir H. Khalifa (b. September 13, 1984-) is an Egyptian scholar specializing in modern English literature. He holds a PhD with first honors (2022) from Kafrelsheikh University. His academic expertise encompasses a range of aspects of English literature, including literary analysis and critical theory. Email: samirhashem@nec.edu.sa & <https://orcid.org/0009-0005-8391-4856>



Ahmad Mohammad Al-Mahamed, born in Syria in 1977, has over twenty years of teaching experience. I hold a PhD in Philosophy, specializing in Arabic Language Curricula, from Mansoura University in Egypt, with honors. I also hold a Master's in Education, specializing in Arabic Language Curricula, from Al Ain University in the UAE, with honors (2020). I also hold a Higher Diploma in Educational Development from Abu Dhabi University in the UAE, with honors (2017). And a Bachelor's in Arts, Arabic Language Department, from Damascus University in Syria (2006). Email: ahmad.almahamed@adu.ac.ae ORCID iD: <https://orcid.org/0009-0001>



Fatma T. G. El-Zaghal is an Assistant Professor of Linguistics, Tanta University, Faculty of Arts, English Department, Tanta, Egypt. Assistant Professor of Phonetics, Tanta Civil University, Faculty of Al-Asun, Egypt. Certified Consultant at Ifad Academy and Head of the English Department at ARTA Journal. She is a Reviewer of scientific researches and a Member of the editorial board. She is a Holder of TESOL certificate from Arizona State University, the United States. Email: fatma_tawakol@art.tanta.edu.eg & <https://orcid.org/0009-0008-2867-9066>

Revisiting Border and Defying Boundaries: Unravelling Echoes of the Partitioned Woman in Geetanjali Shree's *Tomb of Sand*

Bharti Sharma

Department of English, School of Social Sciences and Languages, Vellore Institute of Technology, Vellore, India

Sarika Tyagi*

Department of English, School of Social Sciences and Languages, Vellore Institute of Technology, Vellore, India

Abstract—This paper explores the multifaceted concept of the border, which is geographical, social and psychological, through the perspective of the partitioned woman in Geetanjali Shree's *Tomb of Sand*. It aims to explore how borders serve as repositories of memory and examine diverse perspectives on the history of partition, critically analysing the impact on women and the courage and compassion they exhibited in overcoming these challenges. Furthermore, it investigates the significance of the concept of 'border' through protagonist Ma's experiences and understanding during the partition and in contemporary times. Using insights from border studies and feminist postcolonial theory, the researcher examines how the 1947 Partition continues to shape identity, memory, and femininity. Through Ma's journey across literal and metaphorical borders, *Tomb of Sand* reimagines the border not merely as a site of division and violence but as a space of reclamation, connection, and healing. This study analyses how Shree employs non-linear memory, symbolic spatiality, and intertextual references to Partition narratives to illuminate Ma's suppressed experiences and the courage required to confront them. This paper also argues that Shree's novel offers a transformative framework for understanding the resilience of the partitioned woman, revealing how borders can be crossed, defied, and utilised as a locus of healing. Through this lens, the article contributes to broader discourses on trauma, gendered history, and the enduring human quest to transcend imposed boundaries in both historical and contemporary contexts.

Index Terms—nation, border, partition, trauma

I. INTRODUCTION

"There is no way we can begin to understand what Partition was about unless we look at how people remember it". (Butalia, 2000, p. 10).

Women, borders, and boundaries intersect in profoundly complex ways within the historical and cultural memory of the 1947 Partition of India. The creation and dissolution of borders have long shaped histories, identities, and memories, but few events illustrate their profound human cost as vividly as the Partition of India in 1947. The Partition was not merely the drawing of a political boundary between two nations; it was an act that ruptured identities, families, and communities, leaving deep scars on the social and emotional fabric of the subcontinent. The division has long impacted individuals, but women suffer the most due to being targeted by male aggression for two main reasons: their gender and their faith. The loss of division is catastrophic. As Geetanjali Shree stated during her interview at the Jaipur Literary Festival, "If you think partition happened in 1947 and it's over, you are mistaken. It is very much a part of our lives. We are still a divided family, divided people, and divided hearts" (Singh, 2023).

This momentous political rupture, which divided the subcontinent into India and Pakistan, was not merely a redrawing of territorial lines; it was a violent reconfiguration of lives, communities, and gendered realities. In the midst of mass migrations and communal violence, women's bodies became contested sites upon which notions of honour, nation, and belonging were inscribed. A partition establishes a physical boundary that has a lasting effect on human experiences. In academia, partition has been focused on the "high politics" and "demographic divisions of land" and "facts about the number of migrants" rather than the impact it had on humans. Abduction, rape, forced conversions, and displacements rendered women both symbols and victims of nationalist fervour. Border studies came into existence through the partition literature. According to Sendhardt (2013), the border is not simply geographical, but also social as well as psychological. The study of borders has seen a radical shift in the last 20 years, moving away from the long-held concentration on state borders as unchanging ontological entities with primarily physical qualities. A perspective that emphasises the shifting meaning of borders, their various types and functions, and their social construction has replaced the earlier view of borders as merely geographical demarcations in the recent history of border studies. Wilson states, "borders are expected to be places and spaces of various forms of culture and expression, particularly the nations" (Wilson, 2009, p. 6).

* Corresponding Author. Email: tyagisarika27@gmail.com

The creation of borders thus entailed not only geographical division but also the reinforcement of patriarchal and communal boundaries that redefined women's roles within the emerging nation-states. Revisiting these borders thus involves more than mapping geography; it demands confronting the social and emotional boundaries that partition is imposed and continues to reinforce, and simultaneously navigating immense personal loss and societal stigma. To "defy boundaries" in this context means reclaiming women's voices and experiences that were silenced beneath dominant political narratives. Through the lens of border theory, the exploration of borders becomes an inquiry into memory, identity, and resistance, revealing how women have continually negotiated, challenged, and transcended the limits imposed upon them by history and patriarchy alike. It also shows how women resisted the limitations imposed upon them by reclaiming their agency, bearing witness to their trauma, and challenging the official narratives that sought to silence them.

Geetanjali Shree is a renowned Indian Hindi-language writer who is well known for the exploration of female experience through her work. Geetanjali Shree, through her unusual storytelling methods, beautifully presents the ordinary lives of people. She paints the plot and character through her magnificent use of words and wit. This study involves the exploration of the theme of the reconstruction of the border and the courage of woman in Geetanjali Shree's novel, *Tomb of Sand*. This novel, originally written in Hindi as *Ret Samadhi* (2018), was translated into English by Daisy Rockwell. The book won the International Booker Prize in 2022, making history as the first book translated from an Indian language to win it. *Tomb of Sand* tells the story of Ma, an 80-year-old widow engulfed in sorrow and overwhelming grief following her husband's death. She eventually resumes her relationships with the outside world, and she sets out on a journey to Pakistan, where she spent most of her life before the partition, driven by her inner desire. This novel is characterised by its immersive nature, as it explores the intricate dynamics of human connections and the enduring impact of historical events in contemporary society. It also talks about how women get used to society's limits and stop questioning them. These limitations and boundaries based on unimportant variables grow ingrained in them. It is a meandering, powerful literary journey that constantly veers off the path before finally bringing you back to its central ideas. It is a partitioned tale with magical realism strewn throughout. The border 'land' turns into the border 'sand' as the border transforms into a fluid and kaleidoscopic gateway to a new world.

II. REVIEW OF LITERATURE

Recently, there has been a significant increase in research regarding the impact of partition on women's lives. Some are anthropological, whereas others examine psychological, sociological, political, and economic dimensions of division and post-partition. Jennifer Yusin, in her article "*The Silence of Partition: Borders, Trauma, and Partition History*", enunciates that the history of the 1947 Partition is characterised by borders that were delineated to separate and form the sovereign nations of India and Pakistan, representing both a tangible historical split and an unfathomable source of anguish and pain. She discusses that the Partition-created border becomes the spatial inscription of the meaning and impact of history on the identity of both the collective as a community and as an individual. Shabana Ahmed and Abha Rupendra Pal, in their essay titled "*Partition of India and Women*," primarily provide an account of how women were distressed and tormented both physically and psychologically, viewing them as embodiments of community honour and, secondarily, as sites of community reproduction. Females of both sides had paid the price of independence by renouncing their freedom and femininity. Few scholarly articles exist on Indian border studies, especially with women being the central 'voice'. "*Gender, Memory, Trauma: Women's Novels on the Partition of India*" by Ananya Kabir explains in her essay how women writers like Bapsi Sidwa and Krishna Sobti have written about other women whose lives have been significantly impacted by partition. In this essay, she mainly gives a glimpse of partition as a collective trauma and outlines the different ways in which the event has been recalled and how books penned by women excavate the gendered dimension of such a tremendous event in history. The authors have depicted women as figures of torment and torture during the partition and the catastrophe they have undergone in the novel, such as in Geetanjali Shree's *Tomb of Sand*. Dr Sanket Kumar Jha, in his essay "*A Critique of Geetanjali Shree's Tomb of Sand as a Feminist Metafiction*", enquires into the "understanding of the word 'borders' from a feminine perspective and juxtaposes it against the male-centric interpretation of the term. The essay discusses how Geetanjali Shree uses self-reflection, playfulness and intertextuality, which makes it metafiction. According to Jha, it highlights various issues about contemporary existence, including the roles of gender, class, ethnicity, religion, governance, and others.

III. AN OVERVIEW OF THEORETICAL FRAMEWORK

The origins of border studies can be traced to the field of social geography concerning the U.S.-Mexico border. However, its theories and critiques can be extended to a broader spectrum of boundaries. Theories and critiques have evolved border studies into a more discursive and dynamic discipline, providing several layers of comprehension that enrich social and psychological insight. The source of development is incipient in the seminal border text, Gloria Anzaldúa's *Borderlands/La Frontera: The New Mestiza*, when she claims in the Preface:

The actual physical borderland that I'm dealing with in this book is the Texas-U.S. Southwest/Mexican border. The psychological borderlands, the sexual borderlands and the spiritual borderlands are not particular to the Southwest. The Borderlands are physically present wherever two or more cultures edge each other, where people of different races occupy

the same territory, where under, lower, middle and upper classes touch, and where the space between two individuals shrinks with intimacy (Anzaldúa, 1987, para. 1).

However, at the border of India and Pakistan, the divergence of culture, race, and economic conditions is very minimal. Then the question arises: how is it possible to incorporate the partition and its aftermath within the border studies? Sucharita Sarkar theorises that the Partition established physical boundaries through state-sanctioned violence, which provoked more violence that inflicted psychological scars and established mental barriers. The Partition is inherently connected to the phenomenon of border crossings and the anguish caused by such forced displacement (Sarkar, 2015). According to recent work done on border studies, it is considered that the Palestinian and Indian partitions are the most controversial and traumatic since World War 2. "India's partition created a novel border, separating India from two entities, West and East Pakistan. The new lines did not restore old Mughal jurisdictions... that illustrate fresh cuts" (Wilson & Hastings, 2012, p. 30). These 'fresh cuts' will never age; they still bleed, and they still leave scars on psychics and create unforgettable memories.

Emily Hicks, in her work *Border Writing: The Multidimensional Text*, states that "border writers allow the reader to practice multidimensional perception and non-synchronous memory" (Hicks, 1991, p. 23). Various Partition writers like Manto, Amitav Ghosh, and Bapsi Sidwa use 'multidimensional perception' and 'non-synchronous memory' to give an account of cataclysmic events that happen due to borders created within the nation. Border studies not only focus on the 'divide' it creates but also on the 'uniting'. People have used the term border as a term to portray the 'other' nation as an enemy. Recently, the counter-discourse has been presented by individuals like Butalia (2000), Menon and Bhasin (1998) and others who seek a re-signification of the term 'border' to indicate 'join' rather than 'divide'. As Passi (2001) says, 'borders' are not merely physical, empirical lines or zones that can be frozen on maps and atlases as naturalised 'entities' (Passi, 2001, p. 22). Broader aspects of life and human experience nullify the border. We should look at the border from the perspective of a better world that is more 'unified' and 'healing' rather than with the limited perspective of 'division' and 'hurt'. "The borderlands... are the privileged locus of hope for a better world" (Michaelsen & Johnson, 1997, p. 3). Where the focus of border studies shifts from 'divide' to 'join' and from 'hurt' to 'heal'.

Within post-colonialism, border studies are often read as a legacy of colonialism, particularly in the context of the Partition of India. Bhasin and Menon (1998) argue that the Partition imposed artificial national identities, forcibly separating people based on religion and nationhood. These imposed boundaries generated not only physical violence but also emotional and psychological fractures, especially for women. As Kumar (2001) states, the boundary that was established as a result of the Partition is not only a line on the map; rather, it is recorded into people's minds, bodies, and memories. This observation directly applies to the character of Ma in *Tomb of Sand*, whose body and life were reshaped by the trauma of that division. Her late-life journey becomes a re-inscription of agency onto a landscape that once rendered her powerless. Border studies, thus, offer a multifaceted lens to understand how borders shape and are shaped by history, trauma, gender, and identity. Far from being inert lines, borders become active sites of conflict, memory, and possibility. By interrogating the colonial and patriarchal foundations of modern nation-states, this theoretical framework invites deeper engagement with the voices of those rendered invisible by border regimes while illuminating how memory, culture, and resistance continue to flow across imposed divides. When applied to *Tomb of Sand*, border theory reveals how borders function as both traumatic enclosures and transformative spaces. For Ma, the India-Pakistan border is not only a historical wound but also a site of return, reckoning, and resistance. The perspectives of Gloria Anzaldúa, Sucharita Sarkar, and postcolonial feminist theorists enrich the reading of the novel by challenging the notion of fixed identities and demonstrating that healing and agency often emerge not from avoiding borders but by courageously traversing them. This paper attempts to look at the post-millennial work based on the partition, Geetanjali Shree's *Tomb of Sand* (2018), which reconstructs a border from the metamorphosis of Ma through a courageous journey she took to Pakistan. It explores how Shree reconstructs the border with the help of (1) the 'human experience' and 'voice' of woman through 'non-synchronous memory' and 'multiple perceptions', (2) exploring the border in terms of 'physical' and 'psychological' lenses from woman's perspectives in Partition literature, and (3) analysing how the border adds to the 'healing' of Ma in *Tomb of Sand*.

IV. METHODOLOGY

The researcher used a qualitative research approach to achieve its objectives, focusing on the intersection of border, history, and gender. Both primary and secondary sources were utilised to ensure a comprehensive exploration of the research topic. The novel *Tomb of Sand* by Geetanjali Shree served as the primary text, offering a rich site for thematic analysis, particularly in its nuanced portrayal of the Partition of India and its profound effects on women. In addition to the primary source, a range of secondary materials was gathered to support the research. This included scholarly writings, academic articles, interviews, and various research papers that delve into partition and border studies, with a specific emphasis on women's experiences during this historical period. These secondary sources were meticulously collected from different sources, ensuring access to credible and relevant information. The integration of these diverse sources allowed for a thorough and profound analysis, contributing to a deeper understanding of the complexities related to gender and the historical context of the partition. The methodology is the integration of *border studies*, which offers a critical framework for understanding how physical and metaphorical borders, national, cultural, and gendered, shape individual and collective identities. Border studies enable a deeper engagement with the novel's exploration of transgression,

mobility, and the reconfiguration of boundaries, especially in the context of a female protagonist navigating post-Partition realities.

V. DISCUSSIONS

A. *Border Beyond Lines: The Multidimensional Study of Border - Physical, Social and Psychological Boundaries*

A border is a concept that is a reflection of the idea of a territorial state, and it has been one of the most critical keywords in the history of political geography. In *Tomb of Sand*, Shree intricately narrates the story of an octogenarian widow, Ma, beginning in a middle-class household in Delhi and culminating with an unforeseen twist in the Khyber region of Pakistan, as she seeks to reclaim her identity as Chandraprabha Devi. Through its humorous writing, it demonstrates an in-depth comprehension of the borders and boundaries that are present in every location. A consensus has emerged that boundaries should not be perceived merely as physical lines depicted on maps, atlases, or the terrestrial landscape. These regions serve as sites for political geographers, ethnographers, and anthropologists to delineate border landscapes and the experiences of individuals residing along the border (Passi, 2009). Shree transforms domestic and architectural spaces into powerful metaphors that reflect the psychological and political landscapes of post-Partition India. Among these, walls and doors play a central role in articulating themes of trauma, gendered confinement, resistance, and transformation, particularly through the journey of the protagonist, Ma. Indeed, this work depicts all sorts of border crossings. Be it nation, body, gender, memory, or language. The three parts of the novel, *Ma's Back*, *Sunlight*, and *Back to the Front*, represent three distinct phases of Ma's life. Shree has reconstructed the 'border' through these three phases in the following ways: The first phase of Ma's life, known as "Ma's back", marks the beginning of the narrative. This phase focuses on how the author's extensive use of metaphors, such as walls, doors, and borders, indicates the reconstruction of the border as a social and psychological boundary between family, society, and nation. At the beginning of the story, there is a note that revolves around the border. Shree states, "This particular tale has a border and women who come and go as they please. Once you've got women and a border, a story can write itself" (Shree, 2021, p. 1). Shree reconstructs the idea of a border from the very first line of the novel. The border comes and goes as per the needs of the tale. Shree focuses on the border. In this novel, the border plays a very pivotal role. The border manifests as indelible lines, sometimes visible between nations and sometimes invisible between family members. This opening line introduces significant themes and reflects the author's commitment to giving voice to the often-overlooked stories of women in the context of historical trauma. The border is represented through the metaphor of walls, doors, windows and much more, which indicates social and psychological boundaries between family members and the nation. Shree perfectly blends the border within the narrative through the use of metaphors, multiple perspectives, nonchronological memory, and a stream of consciousness.

In *Tomb of Sand* by Geetanjali Shree, walls and doors serve as powerful metaphors for boundaries, confinement, and transition. These physical elements transform into symbolic spaces that reflect the protagonist Ma's psychological journey and the post-Partition socio-political context. "The wall plays a special role in our story (as do the doors, since you use them to get from one side to another, from here to there, on and on through the centuries, from forever to forever)" (Shree, 2021, p. 16). At the novel's outset, the wall becomes a symbol of Ma's retreat into herself after her husband's death. She sits silently, pressed against the wall, as though merging into it. Shree writes, "She was a bundle, shrinking ever more from moment to moment, sending out a scrambled signal from within her vast quilt that she was still in there somewhere" (Shree, 2021, p. 26). The wall reflects her immobilisation, grief, and the heavy silences around her traumatic past, especially during the Partition. As the narrator remarks, "The wall plays a special role in our story" (Shree, 2021, p. 41), signalling its centrality in representing her inner world. Here, walls frequently represent societal limitations and individual challenges. For instance, the protagonist's reflections on the walls that surround her life—both literally and figuratively—bring to light the fact that she has been captured by the cultural expectations and familial obligations that she is obligated to fulfil. This imagery makes one feel as though they are being suffocated and makes one yearn for freedom. Doors, on the other hand, are symbolic of the possibility of the beginning of something new and the possibility of change. Doors represent transition and awakening. Shree writes, "The door does not dance. Everything passes through it: people, air, colour, laughter, sunlight..." (Shree, 2021, p. 58). This description turns the door into a witness to life's fluidity. When Ma finally moves toward the door and "slips away" (Shree, 2021, p. 95), it marks a symbolic escape from confinement and the beginning of her personal transformation. "The door to Bade's home knows it must remain open no matter what, and there are no constraints on those who enter, in terms of time of arrival, advance notice or knocking before entry. Always free to be and free of charge" (Shree, 2021, p. 27). The act of opening or closing a door in the narrative often signifies pivotal moments in the characters' journeys, such as choices that lead to self-discovery or escape from oppressive circumstances. Together, walls and doors capture the tension between confinement and freedom. While walls represent repression, doors become gateways to healing, movement, and the rewriting of identity.

The last and third part of the novel, entitled 'Back to the Front', starts with the phrase, "Here we are at Wagah, where the tale is drama, and the story is Partition" (Shree, 2021, p. 535). In this section, Shree vividly portrays the group of Partition writers gathered at the Wagah border: Bhisham Sahni, Balwant Singh, Joginder Pal, Manto, Rahi Masoom Raza, Shaani, Intizar Hussain, Krishna Sobti, Khushwant Singh, Ramanand Sagar, Manzoor Ehtesham, Rajinder Singh Bedi. So many, one could go on and on. And crowding around them, actors, bobbling and swirling like acrobats as they read lines from their scripts" (Shree, 2021, p. 535). Through Ma's trip across spatial and temporal boundaries, the writing style deconstructs the idea of borders, both physical and symbolic. Borders, including those between countries, genders, and

identities, are not only crossed but also questioned and deconstructed as the tale unfolds. Passi suggests that, “A boundary does not only exist in the border area, but manifests itself in many institutions such as education, the media, memorial, ceremonies and spectacle. There are effective expressions of narrative linked with boundaries and border conflicts and serve as references to the other” (Passi, 1998, p. 76). In particular, Ma’s relationship with geographical and cultural barriers blurs the line between self and other. As the national borders that divided her after the Partition are reinterpreted as sites of connection rather than division, Ma’s wish to travel to Pakistan once she has recovered from the loss of her husband exemplifies this deconstruction. Ma raises a question to the interrogating officials, “Do you know what a border is ... It is something that surrounds an existence, it is a person’s perimeter..... A border does not enclose; it opens out. It creates a shape.... a border increases recognition. Where two sides meet and both flourish.” She further enlightens them, “A border stops nothing. It is a bridge between two connected parts (Shree, 2021, p. 652).” Here, Shree perfectly portrays the fluidity and flexibility of the border from physical to social, from national to psychological. She asserts that “Do not accept the Border. Do not break yourself into bits with the border. There’s only us. If we don’t accept, this boundary won’t stay” (Shree, 2021, p. 656). Shree defies all the bounds imposed on women in different settings.

B. Mapping Memories: The Border as a Site of Past Experience

Shree presents the border in the novel not just as a physical demarcation between two nations; it represents the fractured memories of the people who lived through the partition. Chauhan states, “Memory is often treated as a debatable concept due to its malleability. It is a fragile concept based on remembering and forgetting. Remembrance can be fragmentary as the events return in flashback. The trauma is ignited through these memories” (Chauhan, 2025, p. 95). In *Tomb of Sand*, the protagonist Ma’s journey back to Pakistan is symbolic of a recovery memory, where the border is no longer a boundary but a place where traumatic and personal past experiences resurface. As Ma navigates her memories, the border itself becomes a map of these memories, showing the distance and gaps that trauma has imposed on her sense of self and identity. Shree describes the border as a site where fractured, lost, and buried memories can flow across, akin to water crossing an open space. This metaphor highlights how the border is a conduit for memory and experience, where what was once divided is now intermingled and fluid.

Ma’s Memories of Partition demonstrate how borders serve as maps, delineating the agonising narratives of expatriation, personal loss, and trauma. Nithiyasri and Pandey state, “Borders are not just lines drawn on the map. They are the mere representation of past events that run through the blood and minds of every survivor who witnessed partition and got displaced from their homeland” (Nithiyasri & Pandey, 2025, p. 3). Memory acts as a cartographic tool, allowing characters to navigate their emotional landscapes. Ma’s memories often take her back to specific locations tied to her experiences during the Partition. These recollections are not merely nostalgic; they evoke feelings of loss and longing, highlighting the emotional weight of historical events. By confronting her memories, she learns to navigate the complexities of her life, transcending the emotional borders that have confined her. These memories are inextricably linked with her identity and the identities of those around her. The notion of the border also serves to map the trauma associated with Partition. Partition was a violent rupture that created irreversible divisions, not only between countries but also within the minds of the people. The traumatic experiences of displacement, violence, loss of family, and the destruction of homes are etched deeply in the memories of the survivors, and these traumatic memories form an invisible border within the psyche of the characters.

Ma’s internal struggle to come to terms with her past is a central theme in the novel, and her memories of Partition are marked by the physical and psychological violence she endured. As Ma reflects on the Partition, she recalls the violent displacements and massacres that tore apart families. The Partition is, for Ma, a painful reminder of the loss of home and identity, as she reflects on the human cost of this artificial division. The violence is implied, haunting her mind and affecting her psychological state: “At the time, everything fell apart the country, the people, and most of all, the people’s minds. What they saw, what they lived through, it left scars, marks they could not rub away” (Shree, 2021, p. 125). This passage speaks to the psychological violence caused by the physical brutality of Partition. The scars Ma refers to are not just physical but emotional, marking a generation of people who lived through atrocities. Despite not being overtly violent, Ma’s psychological withdrawal acts as a form of physical violence. Her immobility symbolises how the physical violence of her past, the massacres, the forced displacements, and the loss of loved ones, have left her incapacitated and unable to move forward. Shree’s depiction of Ma’s life in a state of paralysis also mirrors the paralyzing effects of the partition’s violence: “Ma did not speak. She remained frozen, the years of silence wrapping around her like thick layers of bandages, as though the violence of her youth had made her incapable of doing anything but wait” (Shree, 2021, p. 33).

Shree’s portrayal of Ma’s journey to Pakistan shows how the border becomes a traumatic site that Ma must cross, not only geographically but emotionally. The borders she encounters symbolise the lasting trauma of partition, which continues to resonate in her life and the lives of her family members. “The border was a scar, a reminder of the lives once lived and lost” (Shree, 2021, p. 347). This metaphor emphasises that the physical borders serve as tangible representations of emotional and historical wounds, encapsulating the pain and fragmentation that result from partition. The act of revisiting the site of her past trauma is a form of memory retrieval, where the border is no longer a closed line but an open space of reckoning. The act of crossing the border symbolises her attempt to confront the remnants of that violence, a confrontation with her own past and its ongoing impact on her psyche. “She had never thought she would step into the land that had been her past, her memory, her grief, the violence—now it was time to go there. To face what was lost and what had broken” (Shree, 2021, p. 144).

It shows how Ma's physical journey to Pakistan is an attempt to address the brokenness caused by Partition. The violence of her past is something she must confront directly, as it is inseparable from the pain of her personal history. "It was not the breaking of a boundary that was important, but what lay beyond it" (Shree, 2021, p. 140). This line speaks to the psychological shift that Ma undergoes. The act of crossing boundaries, whether physical, mental, or social, is not merely an escape; it is about the possibilities that lie beyond these limits. Ma's journey is about rediscovering herself and moving toward a future unencumbered by the past. Despite not depicting it in graphic detail, Shree subtly touches on the traumatic experience of physical and sexual violence during India's partition. Instead, Shree portrays the lingering effects of such violence on the psyches of women who lived through it, using Ma's memories as a way to access the border trauma experienced by women during partition.

The scene at the Wagah border serves as a poignant symbol of the complexities surrounding national identity, trauma, and the legacy of Partition. This border, known for its ceremonial flag-lowering event, becomes a multifaceted site of reflection and emotional resonance for the protagonist, Ma. As Ma observes the spectacle at the Wagah border, her emotional response reveals the deep-seated trauma tied to her memories of Partition. The juxtaposition of celebration and sorrow creates a rich emotional landscape. While the crowd cheers and celebrates, Ma is haunted by her past, reflecting on the lives disrupted by the very borders that are being glorified. This contrast evokes a sense of dissonance, highlighting the complexity of identity shaped by trauma. Geetanjali Shree incorporates references to various Partition writers, enriching the narrative and symbolising the complexities of memory, trauma, and identity associated with Partition. The intertextual references to Partition writers create a dialogue between texts, reinforcing the idea that literature transcends geographical boundaries. This aligns with border theory's emphasis on the fluidity and permeability of borders. These literary allusions serve multiple purposes, deepening the thematic exploration of the novel. The inclusion of Partition writers serves as a tribute to the literary tradition that has emerged from the traumatic experiences of the 1947 Partition. By referencing authors like Saadat Hasan Manto, Khushwant Singh, and Bhisham Sahni, Shree acknowledges their contributions to understanding the human cost of Partition. The references to these writers create an intertextual dialogue, allowing *Tomb of Sand* to converse with established narratives of Partition. Each referenced writer adds layers to the narrative, illustrating the multifaceted nature of trauma experienced during Partition. Their varying perspectives highlight different aspects of the human experience, ranging from the brutality of violence to the subtleties of personal loss. Manto's writings often expose the harsh realities of Partition, depicting violence and its psychological effects. His presence in the narrative underscores the rawness of trauma and the difficulty of reconciling with the past. Khushwant Singh's work, which often balances humour with tragedy, reflects the absurdities of human behaviour in the face of crisis. His influence in the novel serves to illustrate resilience and the complexities of human relationships amidst chaos. Throughout the novel, the interactions among women reveal how borders foster a collective memory among those affected by Partition. By sharing their stories, they create a network of remembrance that reinforces their identities and connections to the past. Shree's narrative suggests that while borders can signify division and loss, they can also represent connections to the past. As Ma's understanding of her identity evolves, she begins to see borders as flexible spaces that encompass both memory and past experiences of partition.

C. *Beyond the Barrier: The Border as Locus of Healing*

In Geetanjali Shree's *Tomb of Sand*, borders are not mere physical boundaries demarcated on land but rather depicted as complex and multifaceted structures that embody both despair and healing. "Borders are never to be found only in border areas but are also located in wider social practice/discourse all around societies and increasingly in relation to global space" (Johnson et al., 2011, p. 63). The narrative represents how these boundaries can serve as space for courage, compassion, transformation, and growth, specifically through the journey of the protagonist, Ma. Dasgosh and Hussain state, "Ma's cross-border journey exemplifies this hybrid space, as she moves beyond the rigid boundaries of nation and gender to reclaim and reinterpret the legacy of partition" (Dasgosh & Hussain, 2025, p. 789). Border theory denotes that borders are fluid constructs that influence identities and social connections among people. Therefore, even though there is misery and pain caused due to borders, there is also hope for healing. "You have crossed the border, they reprimand. She chuckles. Anything worth doing transcends borders" (Shree, 2021, p. 17). For the protagonist, Ma, crossing physical borders allows her to regain her identity as Chandraprabha Devi or Chanda, rather than Ma or Amma. Her journey to Pakistan represents a search for self-awareness, which was originally hampered by memories of partition and the resulting anguish. As Ma confronts and resists societal and psychological limitations, her perception of 'self' evolves. This process highlights that identity is a dynamic combination of events and decisions, rather than a static state.

As she navigates her past, she begins to forge a new understanding of herself that transcends the hurt of her experiences. The narrative emphasises her strength in navigating the complexities of borders. Ma's journey reflects the resilience that arises from confronting and transcending the pain associated with displacement. As she learns to embrace her multifaceted identity, she discovers that borders can also be sources of empowerment. The act of revisiting memories associated with borders allows Ma to confront her trauma. By engaging with her past, she transforms her understanding of those experiences. The memories, once painful, begin to serve as sources of strength, enabling her to integrate her history into her present life. Shree suggests that borders are not fixed; they can shift and adapt based on personal journeys. For example, "Like a body engaged in challenging all stereotypes and definitions. A body unrecognizing of the legitimacy of any border. Flowing this way and that" (Shree, 2021, p. 289). With reference to Rosie, Shree explored the fluidity and permeability of the border within bodily space. This fluidity allows for the reimagining of borders as spaces of healing rather than

confinement. Ma's dreams and reflections illustrate this idea, where she envisions crossing boundaries that are not merely physical but also emotional and psychological.

Geetanjali Shree masterfully redefines borders as spaces for healing rather than mere sources of hurt. Through Ma's journey across physical and emotional boundaries, the narrative illustrates the transformative power of memory, community, and resilience. According to Akkamwad, "healing doesn't mean forgetting but rather embracing the past to move forward" (Akkamwad, 2025, p. 174). By embracing her past and forging connections with others, Ma discovers that borders can serve as pathways to empowerment, growth and healing. "At this point, both grandpa-beings spread out their wings..... they crossed some invisible border and lay down on the other side" (Shree, 2021, p. 495). The crossing of the 'invisible border' marks the pathway to healing and growth, overcoming physical and emotional barriers.

VI. CONCLUSION

Geetanjali Shree's *Tomb of Sand (Ret Samadhi)* is about transcending borders, both literal and metaphorical: boundaries about gender, identity, religion, reality, and imagination, as well as arbitrary national borders. Geetanjali Shree sharply and poignantly demonstrates the cataclysmic impacts of partition and unwanted borders, whether they are created between nations, religions, or genders. *Tomb of Sand* tells the tale of an elderly Ma, who defies the conventional Lakshman Rekha of complying with traditions. Through Ma, an octogenarian widow, Shree brings out the impact of partition on women. Shree vividly exhibits the migration and resettlement throughout the novel. Like any other partition novel, Shree exuberantly depicts violence and trauma, which are inextricably linked to the construction of a border through Ma's memories of the past. Furthermore, Shree situates the *Tomb of Sand* at the intersection of history and memory. Shree explores how memory and past events could potentially tell a fictitious tale that transports us on a fictional journey of human migration through time and space. The novel *Tomb of Sand* perfectly captures key aspects of partition literature. It is built upon the quest for self-awareness of an octogenarian widow who transcends all the physical and psychological borders. It flawlessly reconstructed the idea of a border constructed in the text by Geetanjali Shree through recollections of traumatic past events faced by Ma during the Great Partition of India. Ma boldly addresses the lingering anguish of her childhood memories during Partition and reassesses her identity as a mother, daughter, and woman. Geetanjali Shree brings out the critical and timely demonstration of the destructive consequences of borders and boundaries, whether between religions, countries, or genders. Thus, a close reading of Geetanjali Shree's *Tomb of Sand* tells a tale that explores the permeability of borders among nations, societies, religions, genders, languages, oneself, life, and death, surpassing all linguistic barriers, despite all the obstacles and boundaries.

REFERENCES

- [1] Ahmed, S., & Pal, A. R. (2015). Partition of India and women. *Journal of South Asian Literature*, 50(1), 1–15.
- [2] Akkamwad, V. D. (2025). Narrating trauma, resisting erasure: Women's experiences of Partition in *Tomb of Sand*. *International Education and Research Journal*, 11(3), 173–175. <https://doi.org/10.5281/zenodo.15584591>
- [3] Anzaldúa, G. (1987). *Borderlands/La frontera: The new mestiza*. San Francisco: Aunt Lute.
- [4] Butalia, U. (2000). *The other side of silence: Voices from the Partition of India*. Duke University Press.
- [5] Chauhan, A. (2025). 1947 Partition Archive: Contextualising the narrative of trauma and post-memory of the community of survivor. *IIS University Journal of Arts*, 14(1), 94–105.
- [6] Dasghosh, A., & Hussain, S. (2025). Sustaining feminist mnemonics: A study of temporal and spatial ecology of memory in Geetanjali Shree's *The Tomb of Sand*. *Communications on Applied Nonlinear Analysis*, 32(ICMASD), 787–804.
- [7] Hicks, E. (1991). *Border writing: The multidimensional text*. University of Arizona Press.
- [8] Jha, S. K. (2022). A critique of Geetanjali Shree's *Tomb of Sand* as a feminist metafiction. *Literature and Gender Studies*, 11(3), 45–62.
- [9] Johnson, C., Jones, R., Paasi, A., Amoores, L., Mountz, A., Salter, M., & Rumford, C. (2011). Interventions on rethinking "the border" in border studies. *Political Geography*, 30(2), 61–69. <https://doi.org/10.1016/j.polgeo.2011.01.002>
- [10] Kabir, A. (2018). Gender, memory, trauma: Women's novels on the Partition of India. *South Asian Review*, 39(1), 34–50. <https://doi.org/10.1080/02759527.2018.1428946>
- [11] Kumar, K. (2001). *Prejudice and pride: School histories of the freedom struggle in India and Pakistan*. Viking.
- [12] Menon, R., & Bhasin, K. (1998). *Borders and boundaries: Women in India's Partition*. Kali for Women.
- [13] Michaelsen, S., & Johnson, D. E. (Eds.). (1997). *Border theory: The limits of cultural politics*. University of Minnesota Press.
- [14] Paasi, A. (1998). Boundaries as social process: Territoriality in the world of flow. *Geopolitics*, 3(1), 69–88. <https://doi.org/10.1080/14650049808407608>
- [15] Paasi, A. (2001). Europe as a social process and discourse. *European Urban and Regional Studies*, 8(1), 7–28. <https://doi.org/10.1177/096977640100800102>
- [16] Paasi, A. (2009). Bounded spaces in a 'borderless world': Border studies, power and the anatomy of territory. *Journal of Power*, 2(2), 213–234. <https://doi.org/10.1080/17540290903064275>
- [17] Sarkar, S. (2015). The border as resource: Revisiting the Partition. *Cenacle: A Peer-Reviewed Annual Journal of English*, 1(5), 1–9.
- [18] Sekaran, S. N., & Pandey, P. S. (2025). Echoes of Trauma: The Interplay of Bodies and Borders in Geetanjali Shree's *Tomb of Sand*. *Critical South Asian Studies*, 3(1), 53–62. <https://doi.org/10.33182/csas.v3i1.3506>

- [19] Sendhardt, B. (2013). Border types and bordering processes: A theoretical approach to the EU/Polish-Ukrainian border as a multi-dimensional phenomenon. In A. Lechevalier & J. Wielgoths (Eds.), *Borders and border regions in Europe: Changes, challenges and chances* (pp. 21–44). Bielefeld: transcript Verlag. Retrieved May 20, 2025, from <http://www.jstor.org/stable/j.ctv1fxhcq.4>
- [20] Shree, G. (2021). *Tomb of Sand* (D. Rockwell, Trans.). Tilted Axis Press.
- [21] Singh, D. (2023, February 18). “If you think Partition happened in 1947 and it’s over, you are mistaken”: *Tomb of Sand* author Geetanjali Shree. *Moneycontrol*. Retrieved April 15, 2025, from <https://www.moneycontrol.com/news/trends/lifestyle/tomb-of-sand-author-geetanjali-shree-interview-partition-9943621.html>
- [22] Wilson, T. M., & Donnan, H. (Eds.). (2012). *A companion to border studies*. Wiley-Blackwell.
- [23] Wilson, T. M. (2009). Introduction: Experience, memory and locality in the cultures of European borders. *Anthropological Journal of European Cultures*, 18(1), 1–9. Retrieved 22 April, 2025, from <http://www.jstor.org/stable/43234479>
- [24] Yusin, J. (2019). The silence of Partition: Borders, trauma, and Partition history. *Journal of Borderlands Studies*, 34(2), 215–230. <https://doi.org/10.1080/08865655.2017.1402198>



Bharti Sharma is a full-time research scholar in the Department of English, School of Social Science and Languages, Vellore Institute of Technology, Vellore, India. She completed her undergraduate studies at Sharda University, Greater Noida and her postgraduate studies at Gautam Buddha University, Greater Noida, Uttar Pradesh. She is currently pursuing her doctoral research in English literature. Her area of interest is Indian Writing in English and Gender Studies.



Sarika Tyagi is currently a professor and Associate Dean at the School of Social Sciences and Languages at Vellore Institute of Technology, Vellore, India. Her specialisation is in British Literature, Gender Studies & Eco-poetry. She has published numerous papers in Scopus, Non-Scopus and as book chapters. She has also conducted several training programs and delivered lectures on diverse topics. She is also an author and has published a book on poetry and stories; her reviews and articles are published in several newspapers and magazines. For her contribution to academics and literature, she has received honours and awards from across the country. She can be reached at tyagisarika27@gmail.com

Rooted Kinship and Storied Matter: Relational Ontologies and the Ethics of Coexistence in Elif Shafak's *The Island of Missing Trees*

Farhan Ahmad*

Department of English Language and Literature, College of Sciences and Humanities, Prince Sattam Bin Abdulaziz University, Alkharj, Kingdom of Saudi Arabia

Shamsudheen Mannenkuzhiyan

Department of English Language and Literature, College of Sciences and Humanities, Prince Sattam Bin Abdulaziz University, Alkharj, Kingdom of Saudi Arabia

Wahaj Unnisa Warda

Department of English Language and Literature, College of Sciences and Humanities, Prince Sattam Bin Abdulaziz University, Alkharj, Kingdom of Saudi Arabia

Shehnoor Shan

Department of English, Women's College, Aligarh Muslim University, Aligarh, India

Abstract—This article analyzes Elif Shafak's novel *The Island of Missing Trees* (2022) to articulate a stringent relational ontology and an ethics of coexistence that profoundly challenges anthropocentric and ethno-nationalist paradigms. Set against the backdrop of the Cypriot conflict and its diasporic aftermath, the novel foregrounds the interconnectedness between human identity, history, and the non-human world, primarily through the narrative voice of a fig tree. Grounding the analysis in relational philosophy, Indigenous epistemologies, and ecological phenomenology, this study contends that the fig tree functions as a materially embedded, agential, and ethical subject. Its longevity and capacity to archive trauma demonstrate that memory and healing are *intra-active* processes shared across species and generations. By operationalizing concepts such as *intra-action* and the *Honorable Harvest*, the research reveals how Shafak constructs a model of "rooted kinship" that prioritizes interdependence, reciprocity, and a non-human perspective on historical rupture and ecological survival. The novel, therefore, serves as a literary space for reimagining ethical accountability and trans-species coexistence in the context of global ecological and political fragmentation.

Index Terms—relational philosophy, ethics of coexistence, indigenous epistemologies, ecological phenomenology, memory and trauma

I. INTRODUCTION

In recent decades, escalating ecological crises and persistent social upheavals have necessitated a fundamental epistemological and ontological restructuring across academic disciplines (Stockholm International Peace Research Institute, 2021). The prevailing Western frameworks, which typically privilege the autonomous human subject and maintain rigid dualisms such as culture/nature and human/non-human, are increasingly recognized as inadequate for addressing the complex, entangled challenges of the Anthropocene (SIPRI, 2021). This academic reckoning has spurred a literary shift, wherein contemporary fiction often challenges anthropocentric ontologies, exploring modes of ethical existence predicated on deep interconnectedness.

Elif Shafak's novel *The Island of Missing Trees* (2022) exemplifies this critical shift, presenting a worldview rooted in relational ontology that rejects individualism in favor of interdependence between diverse beings. The novel utilizes a split narrative structure and multiple timelines to tell the story of a forbidden romance between Kostas, a Greek Cypriot, and Defne, a Turkish Cypriot, set against the historical backdrop of the ethno-political conflict in Cyprus (1974) and its subsequent diaspora to London. The narrative is critically driven by the unique perspective of a fig tree, which accompanies the surviving family members into exile. The novel thus, explores the intertwined nature of human trauma, migration, and ecological degradation. Crucially, Shafak offers this reflection from the unique perspective of a fig tree. This arboreal narrator acts not merely as a symbolic presence or a passive witness, but as an active agent of ethical and historical reflection, inviting readers to consider pathways for coexistence that transcend dominant anthropocentric paradigms.

* Corresponding Author. Email: f.ahamd@psau.edu.sa

Relational philosophy, which asserts that entities are fundamentally constituted through their relationships rather than existing as isolated, pre-defined units, provides the essential lens for analyzing Shafak's narrative. This view stands in stark contrast to traditional Western metaphysics, which privileges discrete and autonomous subjects (Ingold, 2011; Latour, 2005). This framework is significantly enriched by Indigenous epistemologies, which offer a holistic, reciprocal view of life, and by ecological phenomenology, which emphasizes the sensory and embodied engagement between human consciousness and the natural world (Abram, 1996; Kimmerer, 2013).

While critics acknowledge the novel's focus on interconnectedness, many analyses approach this theme from a metaphorical standpoint, failing to dig into relational philosophy as a stringent theoretical framework. The analytical gap this study addresses is the operationalization and synthesis of high-level relational concepts specifically, utilizing Barad's (2007) *agential realism* and the concept of *intra-action*, and translating Kimmerer's (2013) Indigenous ethics into a precise ethical framework through the *Honorable Harvest*. This synthesis is necessary to fully discover the deeper ontological and ethical implications of interspecies connectedness and to treat the fig tree as an ethical subject capable of moral influence, beyond its role in human narratives.

This study aims to examine *The Island of Missing Trees* through the integrated theoretical frameworks of relational ontology, indigenous epistemologies, and ecological phenomenology; analyze how the novel constructs ethical relationships among humans, animals, plants, and the environment; investigate the representation of trauma, memory, and migration in relation to environmental narratives; and contribute to ongoing debates in the environmental humanities on subjectivity and trans-species coexistence.

By employing a non-human narrator and emphasizing the interconnectedness of different species, Shafak's novel challenges dominant narratives about identity, memory, and agency. It offers a model of coexistence based on reciprocity, shared vulnerability, and ethical responsibility, transcending the boundaries between species and cultures. In this way, *The Island of Missing Trees* aligns with a burgeoning literary and theoretical movement that seeks to redefine human subjectivity and ethics within a broader ecological and cultural context (Bennett, 2010; Tsing, 2015).

The central thesis is that Shafak employs the fig tree as a nexus for synthesizing relational ontology, Indigenous epistemologies, and ecological phenomenology, thereby constructing a radical ethics of coexistence that destabilizes Western concepts of identity, memory, and political agency, positioning arboreal life as integral to cultural and ecological survival.

A. Research Questions

The following research questions guide this study:

1. How does *The Island of Missing Trees* represent relational ontologies between human and non-human actors?
2. In what ways does the novel challenge anthropocentrism through its narrative structure and its use of non-human narrators?
3. How do relational ontologies and ethics of coexistence inform a reading of intergenerational trauma and ecological coexistence in the novel?
4. What implications does the novel offer for reimagining an ethics of care that includes non-human entities and contributes to ecological and social recovery?

B. Research Objectives

1. To examine how the novel *The Island of Missing Trees* addresses the concept of relational ontology, by analysing the interaction between human and non-human characters to highlight the challenge the novel poses to the notion of separating nature from humanity.
2. To investigate how literary narrative strategies, particularly the use of a fig tree as a speaking character, contribute to deconstructing human-centric perspectives and highlighting non-human subjectivities, thus challenging the anthropocentric view prevalent in literary narratives.
3. To explore how the novel's focus on relational ontologies and interspecies ethics contributes to providing a framework for understanding the reciprocal relationship between human memory, history, and environmental continuity.
4. To analyze how the novel fosters a broader ethic of care that includes non-human beings, and proposes a more comprehensive ethical and psychological framework for the recovery of the environment and society.

II. LITERATURE REVIEW

The Island of Missing Trees has garnered significant critical attention for its multilayered examination of historical trauma and interconnectedness, yet a precise, rigorous relational reading remains underexplored.

Scholarly analyses have thoroughly explored the novel through the intertwined lenses of diaspora, migration, and memory. Work by Sumer (2022) highlights the fragmented histories and dislocations experienced by the characters. Specifically, Zehra and Mohsin (2023) argue that Shafak portrays identity as a product of fragmented histories and migrations, demonstrating how the younger generation, particularly Ada, struggles to understand her family's homeland and its landscape, a disconnect that is not merely emotional but also spatial and ecological. This scholarly focus confirms

the necessity of viewing the fig tree as a central agent in re-establishing ties with the natural world that witnessed and bore the trauma.

Further critiques often utilize ecofeminist and postcolonial perspectives, linking the oppression of the female body and the natural world (Fottouh, 2025; DeLoughrey, 2019; Huggan & Tiffin, 2010), while others have affirmed that memory resides in organic forms. Sabbah and Ayuningtyas (2022) explore the novel's Multicultural and Ecocritical elements, confirming a crucial connection between the issues of diaspora and displacement with the health of the natural environment.

Additional studies into the novel's structure support its relational ontology. Sabbah and Ayuningtyas (2022) note that the novel adheres to a cyclical arboreal time, which contrasts with linear human chronology. The structure of the novel, divided into six parts named after the fig tree's life cycle (roots, trunk, branches, etc.), allows the arboreal time to govern the human narrative and write history from a non-anthropocentric viewpoint (Sabbah & Ayuningtyas, 2022). This foregrounding of the fig tree's own temporal structure reinforces the novel's core relational framework (Barad, 2007).

The novel's pioneering narrative form has also been discussed as a crucial ethical strategy. Scholars argue that fictional narrative form provides a unique advantage in representing the subjectivity and interiority of non-human entities, thereby constructing an explicit "ethics of otherness" that facilitates reader empathy and multiplies understanding of the non-human world (Li, 2025).

The engagement with the "arboreal turn" is relevant, with O'Neill (2023) emphasizing the novel's contribution to "Generating new epistemologies and ontologies for the human" (p. 8) and its role in encouraging "New ways of being responsible for nature" (O'Neill, 2023, p. 17). Comparative works (Powers, 2018; Ghosh, 2021) further position Shafak's novel within a movement that centers non-human protagonists and critiques environmental violence. Additionally, extending the political reach of Indigenous epistemologies, scholarly work highlights that the exclusion of Indigenous and local communities' deep spiritual and cultural values toward nature often leads to inadequate policy and unsustainable development (Sangha et al., 2019). This research reinforces the article's ethical framework by validating the importance of incorporating reciprocal, non-utilitarian relationships like the 'Honorable Harvest' into broader conversations about global ecological governance (Sangha et al., 2019).

However, a recurring deficit in existing scholarship is the tendency to recognize interconnectedness metaphorically rather than leveraging stringent theoretical-philosophical frameworks. The failure to utilize the complex tools of *agential realism* (Barad, 2007) or *semiotic cognition* (Kohn, 2013) prevents the full realization of the fig tree's identity as a true ethical subject capable of agency and moral influence. The present study intervenes by providing this missing theoretical rigor, moving the analysis from thematic observation to ontological proof, strictly within the parameters of relational philosophy, ecological phenomenology, and Indigenous ethics.

III. THEORETICAL FRAMEWORK: SYNTHESIZING ONTOLOGY, PERCEPTION, AND ETHICS

To rigorously analyze Shafak's novel, it is imperative to establish a robust theoretical foundation drawn from relational thought that validates the fig tree's active narrative and ethical role. The three core theoretical domains below are mutually dependent: Relational Ontology defines what exists (entanglement), Ecological Phenomenology defines how that existence is perceived and communicated (cognition), and Indigenous Epistemology defines how we should act within that existence (ethics).

A. Relational Ontology: Agential Realism and Intra-Action

Relational ontology fundamentally posits that identity and being are not inherent properties of isolated entities, but rather emerge dynamically through mutual relationships and co-constitution (Barad, 2007; Latour, 2005). This approach emphasizes interconnectedness and dynamic becoming (Coccia, 2020).

Barad's (2007) agential realism asserts that reality is differentiated by internal interactions, thereby challenging the conventional notion of interaction (which assumes separate agents). Instead, Barad introduces the concept of "Intra-action" (Barad, 2007, p. 141). Intra-action recognizes agency as a "Dynamism of forces" in which all entities are "Constantly exchanging and diffracting, influencing and working inseparably" (Barad, 2007, pp. 133, 141). Entities do not precede their relations; they emerge from them. This framework is essential for establishing the fig tree not as a passive backdrop, but as a material, co-constitutive force in human history and trauma.

B. Ecological Phenomenology and Non-Human Cognition

Ecological phenomenology explains how subjectivity is tied to sensory engagement with the natural world. Abram's (1996) work demonstrates that human consciousness and a sense of belonging arise through profound, embodied interactions with the immediate ecological context. This ethical framework emphasizes that human perception is fundamentally inseparable from the environment (Abram, 1996).

This view is powerfully augmented by Kohn's (2013) work, which argues that non-human life forms, including forests, engage in semiotic processes; interpreting and responding to signs thereby, expanding the scope of cognition and subjectivity beyond human language. Kohn's concept of an "ecology of selves" is crucial, as it suggests that individuals, both human and non-human, are defined through their relationships via sign interpretation. Shafak's fig tree operates within this domain, providing the necessary evidence for non-human cognition to ground the novel's ethical claims.

C. Indigenous Epistemologies and the Ethics of Kinship

The relational and phenomenological insights establish that entanglement exists; Indigenous epistemologies provide the prescriptive ethical foundation for how to honor it. Indigenous scholars emphasize a holistic, relational view of life, promoting an ethics founded on respect, responsibility, and profound interconnectedness between human beings and the land (Kimmerer, 2013; Simpson, 2017). Central to this ethical framework is Kimmerer's (2013) concept of the 'Honorable Harvest' a sophisticated ethical protocol based on mutualistic reciprocity. The principle demands asking permission before taking, only taking what is needed, and offering a gift in return, ensuring that human actions are based on kinship and respect for the agency of non-human beings (Kimmerer, 2013). This perspective inherently challenges the colonial and anthropocentric structures that view nature as inert property (Plumwood, 2002; Watts, 2013).

The tragic backdrop of ethno-nationalist conflict in Cyprus, characterized by colonial partitions and ecological destruction, fundamentally represents a profound failure of the Honorable Harvest a "dishonorable harvest." Kostas's observation that the land "Was not the verdant paradise he remembered. The absence of trees was a powerful rebuke to the dreadful mistakes of the past" (Shafak, 2022, p. 198) materializes this ethical failure. The environmental degradation, symbolized by the missing trees, is not merely an ecological phenomenon, but also a moral one. The trees become symbols of presence and absence, of memory and loss. Their disappearance reflects the historical violence inflicted upon both people and the land. Shafak's novel highlights this rupture, offering a perspective that encourages reflection on the ethical dimensions of environmental degradation and calls for reconciliation with nature (Kimmerer, 2013).

This relational perspective compels an ethical expansion, extending moral concern beyond the human community to encompass the non-human world (Plumwood, 2002). Haraway's (2008, 2016) work, which champions the practice of becoming-with and the call to "make kin" across species, provides a foundational argument for a relational mode of existence based on mutual care and interdependence.

The fig tree's suffering embodies this shared moral significance. When the tree reflects that "It was not just the people who had suffered; the land had, too. And the trees. And the animals. Everyone suffered" (Shafak, 2022, p. 190), the narrative aligns with an environmental ethic that recognizes shared vulnerability. By granting the fig tree moral influence and narrative authority, Shafak performs a literary extension of ethical philosophy, contributing directly to debates concerning ethical accountability for non-human entities (Aspoy, 2024; Levinas, 1991).

IV. METHODOLOGY

This study adopts a qualitative and interpretive methodology, drawing on a literary analysis grounded in ecocriticism and relational philosophy. The primary approach is a close reading of Elif Shafak's novel *The Island of Missing Trees* (2022), paying particular attention to narrative structure, symbolism, character development, and the use of non-human narratives, most notably the fig tree, a key voice in the novel. The theoretical framework integrates the insights from relational ontology, Indigenous epistemologies, and ecological phenomenology to explore the interconnected webs of existence represented by the novel, in which plants, animals, humans, and landscapes exist in dynamic interdependence. A contextual analysis is applied by reading the novel against the backdrop of the Cyprus conflict and diasporic identity, enriching the interpretation of political and environmental concerns. This interdisciplinary approach enables a nuanced reading of the novel that highlights the ethics of coexistence and the importance of supra-human perspectives in literary narrative.

V. INTEGRATED ANALYSIS

Shafak's *The Island of Missing Trees* (2022) interrogates conventional concepts of ownership and identity, encouraging analysis through the framework of relational philosophy. The novel examines life, death, and the interconnectedness of living organisms. By featuring a fig tree that articulates perspectives on humanity, the narrative foregrounds the relationship between humans and the natural world. Employing relational ontology and ecological phenomenology, Shafak contrasts human suffering with ecological distress resulting from anthropogenic activities, highlighting the reciprocal effects of these interactions. The fig tree, as a narrative device, advances a relational ontology that situates humans and nature within a network of interdependent relationships. In the context of environmental crises, this approach facilitates a nuanced understanding of emergent forms of nature and underscores the imperative for ethical coexistence.

This section integrates the application of the synthesized theoretical framework to the novel's text with a comprehensive discussion of the findings and their broader implications for trans-species coexistence and societal critique.

A. The Intra-Active Materiality of Storied Memory

In *The Island of Missing Trees*, Shafak intertwines human lives with nature, thus creating a concept of interconnected materiality, where memory is not confined to the human mind or narrative, but is embedded in living entities, particularly the fig tree, and emerges through them. Drawing on Barad's (2007) theory of *intra-action*, which emphasizes the interconnectedness of all things, the novel demonstrates that memory is not confined to human consciousness, but is distributed throughout all entities, both human and non-human. The fig tree, transplanted from Cyprus to London, serves as a tangible symbol with profound meaning, reflecting both personal and collective suffering. As O'Neill (2023) explains,

the tree is not merely a symbol, but a “character” that actively participates in the narrative, blurring the boundaries between the animate and inanimate, and between what speaks and what remains silent.

The fig tree functions as a dynamic, storied matter and a “multi systemic archive of trauma and regeneration.” Its move from Cyprus to London reflects the characters’ diasporic journeys and speaks to relational ontology; the view that beings do not exist independently, but are defined through their relationships. “A tree always knows that it is linked to endless life forms and that its existence is not isolated happenstance but intrinsic to a wider community” (Shafak, 2022, p. 100). The tree’s memory processes provide compelling textual evidence for relational ontology (Barad, 2007) “Across history we have been a refuge to a great many. A sanctuary not only for mortal humans, but also for gods and goddesses” (Shafak, 2022, p. 178).

Memory is a material-discursive phenomenon, not exclusive to the human brain. The memory of conflict is physically constituted *through* its organic engagement with the earth, defining intra-action. The tree’s roots are literally entangled with the “Sinews of history, the ruins of war nobody came to win, the bones of the missing” (Shafak, 2022, p. 211). This material entanglement makes the fig tree a co-constitutor of this history (Barad, 2007). This ontological foundation immediately informs the environmental ethic (Plumwood, 2002). Because the fig tree’s memory and being are intra-actively constituted by human trauma, its suffering is not peripheral, but central: “Humans were not the only one that suffered. So, did we trees - and animals too, experienced hardship and pain as their habitats came to disappear. It never meant anything to anyone, what happened to us” (Shafak, 2022, p. 190). This shared vulnerability mandates a shared ethical responsibility, reinforcing the necessity of viewing human suffering and ecological degradation as a singular moral consequence (Haraway, 2008).

Furthermore, the novel’s narrative structure challenges anthropocentric perspectives by giving the fig tree a central role in the narrative. By attributing human characteristics, the fig tree as a narrative strategy emphasises idea that human notions of identity and belonging are intimately connected to the natural environment. This approach allows Shafak to explore ethical issues related to care and coexistence. Laheg (2024) suggests that the tree functions as a “moral witness,” documenting the silences and absences in both human and natural history. By listening to and connecting with the tree’s stories, the novel’s protagonists and, by extension, the readers are encouraged to adopt an ethic of care that transcend purely human concerns.

The interrelationship between the human and non-human worlds is further illustrated by the novel’s metaphorical portrayal of the vegetal process of implanting and digging up the tree, a beautifully created figurative image that both apprises of the adversarial reminiscing on the past and suggests compromise and recoupling with one’s roots no matter how tangled they are.

In this sense, the storied memory of the novel is fundamentally relational: it emerges from the interaction between human beings, plants, history, and place. The mobility, rootedness, and scars of the fig tree are not mere metaphors, but rather constitute a central element of the novel’s ecological and political narrative. As O’Neill (2023) points out, the perspective of the tree transcends temporal and spatial distances, uniting in a single narrative voice the memories of war, exile, and environmental degradation. In *The Island of Missing Trees*, memory is not an abstract or purely mental function; it is material, embodied, and constructed collectively by different species and nature itself.

B. Non-Human Cognition and the Ethical Critique

In her novel *The Island of Missing Trees*, Elif Shafak offers an ethical critique of anthropocentrism by portraying non-human cognition through the voice of the novel’s protagonist, the fig tree. This challenges conventional notions of intelligence, sentience, and moral value. The tree is not merely a passive object of human observation; it is a thinking, remembering, feeling, and narrating being (a non-human cognition) that participates in moral testimony. For example, the tree states, “A tree’s rings do not only reveal its age, but also the traumas it has endured... carved deep in each circle, is a near-death experience, an unhealed scar” (Shafak, 2022, p. 45). This suggests that the tree’s memory is inherent in its structure, and that it processes trauma in a way like humans. Similarly, when Kostas buries the tree to protect it from the harsh London winter, this act is viewed as a moral responsibility towards a living being in need of protection. “A tree is a memory keeper... Tangled beneath our roots, hidden inside our trunks, are the sinews of history, the rains of wars nobody came to win, the bones of the missing” (Shafak, 2022, p. 211). This implies that the tree carries a collective memory across generations, possessing a consciousness that resonates with moral testimony.

The fig tree’s narrative authority is validated through ecological phenomenology (Kohn, 2013), which asserts non-human cognition. The tree describes its complex communication network: “We trees communicate all the time, sending chemical signals through the air and across our shared mycorrhizal networks” (Shafak, 2022, p. 99). This affirmation of a distributed, functional intelligence challenges the humanist requirement of a central brain for subjectivity. The tree immediately utilizes this cognitive authority to launch an ethical critique against anthropocentrism. The tree notes: “Human do not want to ascertain whether we may be capable of volition, altruism, and kinship. They find it easier to assume that trees can only experience the most rudimentary existence” (Shafak, 2022, p. 44). This act of philosophical evaluation demonstrates that the non-human world possesses the capacity to interpret and morally judge human actions, aligning with the philosophical extension of ethics to the “Other” (Levinas, 1991).

C. Reciprocity, Healing, and the Honorable Harvest

The culmination of the ontological and perceptual findings is the necessary shift to an ethical practice. The tragic failure of the dishonorable harvest the colonial and nationalist exploitation of the land is redeemed through the relationship between Kostas, Defne, and the fig tree, which embodies the Indigenous ethic of Honorable Harvest (Kimmerer, 2013). Kostas's labor and care for the transplanted sapling in London giving resources, knowledge, and protection is reciprocated by the tree, which offers shade, memory, and spiritual anchoring (Shafak, 2022). This is an active, mutualistic practice of *becoming-with* (Haraway, 2008). The healing of the daughter, Ada, depends entirely on this trans-species relationship. Her emotional stabilization occurs when she leans against the tree, feeling her breath synchronize with its "silent rhythm" (Shafak, 2022), demonstrating that reconnection with her fractured human history requires re-establishing physical and emotional rooted kinship with the non-human entity that holds that history. The tree thus serves as a model for how the ontological truth of entanglement must be matched by an ethic of reciprocity to achieve reconciliation.

Shafak's narrative powerfully deploys these integrated findings to offer a profound ethical and societal intervention. The novel uses the fig tree's relational existence as a direct blueprint for condemning human division and promoting ethical coexistence. The tree's commentary "Under and above the ground, we trees communicate all the time. We share not only water and nutrients, but also essential information. Although we have to compete for resources sometimes, we are good at protecting and supporting each other" (Shafak, 2022, p. 99), is a direct ecocentric critique of exclusionary societal systems. This statement is substantiated by the finding of intra-action (Section 6.1) because their existence is materially co-constituted (Barad, 2007), arboreal life is structurally dependent on cooperation, a necessity that transcends the arbitrary boundaries of human identity.

The novel rejects linear, anthropocentric time, adopting instead the cyclical, rhizomatic pattern of the tree (Deleuze & Guattari, 1978). This structural choice reinforces the Indigenous epistemological finding that the dishonorable harvest is unsustainable. The fig tree insists on a paradigm of *interbeing*, where non-human moral influence (Aspoy, 2024; Plumwood, 2002) guides societal and ecological reorganization.

Besides, the fig tree's role as a mobile memory archive reflects the necessity of multidirectional memory (Rothberg, 2009) for diaspora healing. The tree, carrying the "sun and the dust of the Mediterranean inside me" (Shafak, 2022, p. 33), materially embodies the sensory entanglement with place, a concept central to ecological phenomenology (Abram, 1996). The healing of Ada requires recognizing that her family's past is not abstract history, but is physically archived in the tree, underscoring that human healing is inseparable from ecological restoration. By granting the fig tree the status of an ethical subject, Shafak envisages the reader to acknowledge that reconciliation from historical violence is achievable only when the non-human entities that witnessed and bore the trauma are recognized as integral partners in the process of regeneration. The novel confirms that ecological and cultural survival requires recognizing the non-human world, in its materiality and agency, as an equal partner in the processes of healing and history-making.

This fictional representation of arboreal life and its human-like agency "Generates new epistemologies and ontologies for the human" (O'Neill, 2023, p. 8). The convergence between science and sensuality in the novel suggests new forms of connections. By humanising the fig tree, Shafak breaks down the barrier between humans and nature and forces us to appreciate them for what they are and what they give. She proposes that treating nature as an object creates a barrier between us, exonerating ourselves of moral accountability and encouraging exploitation. The description of the interrelationship of the fig tree with the arboreal ecosystem as depicted in the novel informs about Shafak's arboreality.

A tree always knows that it is linked to endless life forms – from honey fungus, the largest living thing, down to the smallest bacteria and archaea – and that its existence is not an isolated happenstance but intrinsic to a wider community. Even trees of different species show solidarity with one another regardless of their difference, which is more than you can say for so many humans. (Shafak, 2022, p. 100)

Shafak's emphasis on arboreal matter, especially her treatment of the fig tree, shows her knowledge about the trees and a deep appreciation for nature. Her humanization of the fig tree expands human ways of knowing and understand the natural world. Her projection of the fig tree as, familiar, friendly, and connected to lively beings challenges the traditional human/plant hierarchy in favour of a relational one. This arboreal association and exchange according to O'Neill (2023) "Highlights the novel's significant contribution to the imaginative work needed in the Anthropocene to find new ways of being responsible for nature" (p. 17). Since the relationship between humans and nature is reciprocal, any change made by one entity can affect the other. Shafak explores this interrelationship between humans and nature on different levels such as spiritual, economic, health, food, and arboriculture.

Through these descriptions, Shafak raises an ethical critique. She compels us to reconsider the idea that moral values and cognitive abilities are exclusively human attributes. Shafak's portrayal of the fig tree's capacity to perceive and interact with its environment has ethical implications: if trees can remember, suffer, and possess moral value, then human actions that ignore, exploit, or harm them are immoral. For example, human violence in Cyprus (bombings, territorial division, displacement) demonstrates that it not only affects people but also harms the fig tree, thus underscoring the novel's perspective on our ethical responsibility towards nature. The novel suggests that environmental damage is also moral damage; harm to nature is harm to memory and human bonds. As the study, *Branching Paths of Pain: Trauma and Ecocriticism in Elif Shafak's The Island of Missing Trees* (2024) concludes, "The fig tree symbolizes memory, resilience, and healing, and serves as a witness to both human and environmental suffering" (Elgamal, 2024, p. 37). In short, the concept of non-human sentience that Shafak presents in the novel encourages a redefinition of ethics, where humanity is not the sole moral agent, but part of an interconnected ecosystem, whose cognitive and moral capacities deserve respect

and care. Through her unique narrative technique, Shafak challenges us to reconsider our relationship with the natural world and the impact of our action on it.

VI. CONCLUSION

This study demonstrates that Elif Shafak's *The Island of Missing Trees* (2022) is a theoretically sophisticated literary text that moves far beyond metaphor to articulate a rigorous philosophy of coexistence. Through the narrative of the fig tree, Shafak successfully synthesizes complex theoretical frameworks: Barad's (2007) relational ontology and the concept of *intra-action* establish the tree as a material co-constitutor of history; Kohn's (2013) non-human semiotics validate the tree's distributed cognition and narrative authority; and Kimmerer's (2013) Indigenous ethics, expressed through the framework of the *Honorable Harvest*, prescribe the necessary moral foundation for human-nonhuman reciprocity.

The integrated analysis confirms that the fig tree functions as a fully realized non-human ethical subject and a trans-species memory archive. The novel effectively critiques anthropocentric systems, showing that nationalist conflict and ecological degradation are two sides of the same dishonorable harvest that embodies a failure of relational responsibility. By centering the perspective of the fig tree, Shafak validates the concept of 'storied matter' asserting that the environment is not a passive backdrop but an active repository of the history and memory whose agency must be recognized. Conversely, the arboreal solidarity and the practice of mutual care between the tree and the human characters offer a model of "rooted kinship" that provides a blueprint for ecological and societal reconciliation.

The novel serves as a powerful testament to the idea that ecological and cultural survival requires recognizing the non-human world, in its materiality and agency, as an equal partner in the processes of healing and history-making. Future research could extend the application of *agential realism* to analyze other narratives of post-conflict reconstruction, specifically examining how non-human elements resist fragmentation in diasporic literature.

ACKNOWLEDGEMENTS

This study is supported via funding from Prince Sattam bin Abdulaziz University, project number (PSAU/2025/R/1447).

REFERENCES

- [1] Abram, D. (1996). *The Spell of the Sensuous: Perception and Language in A More Than Human World*. Pantheon Books.
- [2] Aspoy, H. (2024). Forest Science Between Human and Nonhuman Agency. *Trees, Forests, and People*, 15, 1-8. <https://doi.org/10.1016/j.tfp.2023.100485>
- [3] Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.
- [4] Bennett, J. (2010). *Vibrant matter: A political ecology of things*. Duke University Press.
- [5] Coccia, E. (2020). *The life of plants: A metaphysics of mixture*. Polity Press.
- [6] DeLoughrey, E. (2019). *Allegories of the Anthropocene*. Duke University Press.
- [7] Elgamal, A. M. A. (2024). Branching Paths of Pain: Trauma and Ecocriticism in Elif Shafak's *The Island of Missing Trees*. *Egyptian Journal of English Language and Literature Studies*, 13(1), 37-60. <https://doi.org/10.21608/ejels.2024.413598>
- [8] Fottouh, N. (2025). An eco-feminist depiction of Elif Shafak's *The Island of the Missing Trees*. *Journal of the Faculty of Arts-Fayoum University*, 17(1), 606-639. <https://doi.org/10.21608/jfafu.2025.362994.2188>
- [9] Ghosh, A. (2021). *The nutmeg's curse: Parables for a planet in crisis*. University of Chicago Press.
- [10] Gibson-Reinemer, D. K., Sheldon, K. S., & Rahel, F. J. (2015). Climate Change Creates Rapid Species Turnover in Montane Communities. *Ecology and Evolution*, 5(12), 2340-2347. <https://doi.org/10.1002/ece3.1518>
- [11] Guattari, F., & Deleuze, G. (1978). *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press.
- [12] Haraway, D. J. (2008). *When species meet*. University of Minnesota Press.
- [13] Haraway, D. (2016). *Staying With the Trouble*. Duke University Press.
- [14] Huggan, G., & Tiffin, H. (2010). *Postcolonial ecocriticism: Literature, animals, environment*. Routledge.
- [15] Ingold, T. (2011). *Being alive: Essays on movement, knowledge and description*. Routledge.
- [16] Kimmerer, R. W. (2013). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge and the teachings of plants*. Milkweed Editions.
- [17] Kohn, E. (2013). *How forests think: Toward an anthropology beyond the human*. University of California Press.
- [18] Laheg, I. (2024). Identity crisis as intergenerational trauma progeny in Shafak's *The Island of Missing Trees*. *New Literaria: An International Journal of Interdisciplinary Studies in Humanities*, 5(1), 1-7. <https://newliteraria.com/articles/v05i1/abs-v05i1-01.pdf>. Retrieved 29 November, 2025, 12:31:04.
- [19] Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- [20] Levinas, E. (1991). *Totality and infinity: An essay on exteriority* (A. Lingis, Trans.). Duquesne University Press. (Original work published 1969)
- [21] Li, Yafei. (2025). Encountering the Non-Human with Narrative Form: J. M. Coetzee's *The Lives of Animals*. *Kritika Kultura*: 47(10), 193-206. <https://doi.org/10.13185/1656-152X.2186>
- [22] O'Neill, S. (2023). Arborealities, or Making Trees Matter in Elif Shafak's *The Island of Missing Trees*. *ISLE: Interdisciplinary Studies in Literature and Environment*, 31(4), 796-816. <https://doi.org/10.1093/isle/isad040>
- [23] Powers, R. (2018). *The overstory*. W. W. Norton & Company.

- [24] Plumwood, V. (2002). *Environmental culture: The ecological crisis of reason*. Routledge.
- [25] Rothberg, M. (2009). *Multidirectional memory: Remembering the Holocaust in the age of decolonization*. Stanford University Press.
- [26] Stockholm International Peace Research Institute. (2021). *Anthropocene (in)securities: Reflections on Collective Survival 50 years after the Stockholm Conference* (E. Lövbrand & M. Möbjörk, Eds.). Oxford University Press.
- [27] Sabbah, S., & Ayuningtyas, P. (2022). The issues of diaspora and displacement in Elif Shafak's *The Island of Missing Trees*. *Rainbow: Journal of Literature, Linguistics and Culture Studies*, 11(2), 62-69. <https://doi.org/10.15294/rainbow.v11i2.58682>
- [28] Sangha, K. K., Russell-Smith, J., & Costanza, R. (2019). Mainstreaming indigenous and local communities' connections with nature for policy decision making. *Global Ecology and Conservation*, 19, 1-13. <https://doi.org/10.1016/j.gecco.2019.e00668>
- [29] Shafak, E. (2022). *The Island of Missing Trees*. Penguin Books.
- [30] Simpson, L. B. (2017). *As we have always done: Indigenous freedom through radical resistance*. University of Minnesota Press.
- [31] Sumer, M. (2022). Memory and nature in Elif Shafak's *The Island of Missing Trees*. *Journal of Literary Studies*, 38(4), 421-435.
- [32] Tsing, A. L. (2015). *The mushroom at the end of the world: On the possibility of life in capitalist ruins*. Princeton University Press.
- [33] Watts, V. (2013). Indigenous place-thought and agency amongst humans and non-humans (First Woman and Sky Woman go on a European world tour!). *Decolonization: Indigeneity, Education & Society*, 2(1), 20-34. <https://jps.library.utoronto.ca/index.php/des/article/view/19145>. Retrieved 27 November, 2025, 9:40:04.
- [34] Zehra, S., & Mohsin, L. A. (2023). Rootless Identity in Shafak's *The Island of Missing Trees*: A Postcolonial Perspective. *Journal of Policy Research*, 9(1), 395-401 <https://doi.org/10.5281/zenodo.8000551>

Farhan Ahmad is currently working as an Assistant Professor of English at the Department of English Language and Literature, College of Science and Humanities, Prince Sattam bin Abdulaziz University, Alkharj, Kingdom of Saudi Arabia. He is an active researcher and published a number of articles in Scopus and Web of Science. His research interests include performance studies, Modern European drama, Cultural studies, Gender studies and ESL/EFL pedagogy.

Shamsudheen MK is an accomplished academic with extensive teaching and research experience across various universities and colleges in India and abroad. He currently serves as an Assistant Professor of English at Prince Sattam University, KSA.

Wahaj Unnisa Warda has been working in the Department of English Language and Literature, Prince Sattam Bin Abdulaziz University, Saudi Arabia for the past 10 years. She has presented and published papers at various international conferences and journals in London, India, Australia, and Saudi Arabia. She is a member of various national and international organizations, as well as a reviewer for international journals. Her areas of interest are psychoanalysis, classical mythology, and detective fiction of the 19th century.

Shehnoor Shan is an Assistant Professor of English at Women's College, Aligarh Muslim University, Aligarh, India. She is the recipient of Full Bright Scholarship, and has worked as a Full Bright Fellow and Cultural Ambassador at the University of Texas at Austin, USA. She also works as an English Editorial Fellow at the Foundation for Evidence Based Development in India.

Language Change in Social Media and Its Impact on the Linguistic Identity of Indonesia's Younger Generation: A Sociolinguistic Study

M. Ridha Ajam

English Literature Study Program, Faculty of Cultural Sciences, Khairun University, Indonesia

Mujahid Taha*

National Research and Innovation Agency of Republic Indonesia (BRIN), Indonesia

Andi Sumar-Karman

Social Anthropology Study Program, Khairun University, Indonesia

Abstract—Current communication patterns in Indonesia have changed due to the rapid development of digital technology, especially among teenagers. Social media sites such as Instagram, Facebook, and YouTube are not only a means of sharing information but also interactive spaces for creating and expressing individual and group identities. Using a qualitative descriptive approach and literature review, this study aims to analyze the types of language transformation in social media and their impact on the linguistic identity of Indonesian teenagers. Data was collected from several digital platforms, and the analysis revealed that digital communication fosters the emergence of new linguistic forms, such as slang, acronyms, code-mixing, and abbreviations. Therefore, a balanced approach in language policy and education is necessary to support innovation and the maintenance of linguistic and cultural diversity in the digital age.

Index Terms—sociolinguistics, social media, adolescent identity, code-mixing, language change

I. INTRODUCTION

Especially among the younger generation, the development of digital technology has shaped contemporary communication patterns. Social media platforms such as Instagram, Facebook, and YouTube not only serve as tools for sharing information, but also as interactive social spaces that shape and display individual identities. The use of language on social media serves not only as a form of personal expression, but also as a tool for building identity, conveying information, and interacting effectively with audiences, especially the younger generation. In addition, these platforms facilitate the creation of dynamic content that is relevant to trends, so the language used tends to be informal and reflects developments in popular culture (Arévalo-Salinas et al., 2025). The younger generation uses these platforms to express themselves, build social networks, and negotiate their identities. Communication is now dynamic, contextual, symbolic, and one-way or statistical. This communication pattern is accompanied by a transformation of language. On social media, the younger generation creates new forms of language, including abbreviations such as XYI, gws, acronyms, code-mixing, code-switching, and the use of visual symbols such as emojis to convey meaning. Language functions as a context for fast and informal conversation that is more adaptively used. This is in line with the media and social needs of its users, reflecting its dynamic and adaptive nature (Biagi & Suryadmaja, 2025).

The use of language on social media reflects the social identity of the younger generation. It displays individual expression but also indicates group affiliation, social status, and cultural aspirations. The use of Indonesian and English, for example, is contemporary, global, and urban in style. Language here functions as a symbol of social identity, which is consciously constructed through digital interaction. To understand how language shapes and reconstructs identity in the digital realm, a sociolinguistic approach is very relevant. This is because sociolinguistics considers language to be a social and cultural product that cannot be separated from its context. In the realm of social media, language variation reflects linguistic differences and complex social dynamics, including power relations, ideology, and changing cultural values (Asdah & Safitri, 2025).

One aspect of the digital linguistic phenomenon is the prevalence of slang. Slang represents closeness, modernity, and solidarity among the younger generation. The use of acronyms such as *gabut*, *baper*, or even the word *ngeri* is not only to convey meaning, but also to build a sense of solidarity among community members. Slang can function as an exclusive mechanism that limits understanding to certain groups (Iswatiningsih et al., 2021). In addition, the language used is usually in the form of abbreviations, shortened words, and non-standard variations of Indonesian that are adapted to the needs of users, especially teenagers. The use of language on social media often leads to bilingualism or

* Corresponding Author.

multilingualism (Mardikantoro et al., 2023). This raises concerns that it may reduce the use of formal Indonesian in digital media. Social media, which is rich in informal language use, is considered to be damaging the structure of the national language. Language variation is a natural process in multilingual societies, but from a sociolinguistic perspective, it is assumed that adaptation to context is not a deviation (Putri & Sofyaningrum, 2024).

In addition to slang, code-mixing and code-switching occur on social media. The use of code-mixed language on social media is a sociolinguistic phenomenon in which Generation Z uses English and Indonesian simultaneously when communicating on social media platforms. This phenomenon is influenced by the development of information and communication technology (ICT), which accelerates interaction and self-expression through social media. Code-mixing is done for various reasons, such as to facilitate communication, express identity, demonstrate a higher level of knowledge, and make people more familiar with others (Tarihoran et al., 2022). Young people often mix Indonesian, English, and regional languages in a single post or conversation. The use of code-switching not only reflects language proficiency, but also serves as a communication strategy to build identity, adapt to the audience, or emphasize certain meanings (Sitanggang et al., 2025). In digital communication, meaning is also conveyed through nonverbal forms such as emojis, GIFs, and other visual symbols. Conveying emotions, for example, functions as the use of intonation or emphasis. Social media has become multimodal. Therefore, current linguistic studies are not only focused on text, but also include aspects of linguistic practice (Kittiveja et al., 2025).

Social media has become an arena for negotiable linguistic ideology. It reflects the tension between preserving the national language and openness to globalization, discourse on standard language and slang. The younger generation is actively involved in shaping the values, norms, and direction of the language they use every day. Given this complexity, it is important to collect real data from social media. Digital linguistic research can help map linguistic practices and their impact on social identity, such as discourse analysis or corpus linguistics. It is important to support a language that is inclusive, adaptive, and not trapped in linguistic conservatism alone (Natsir et al., 2023). A language is needed that is capable of connecting national language norms with the digital linguistic reality of the younger generation. The government and educational institutions must actively promote the use of Indonesian without hindering the linguistic creativity that is developing on social media. Language must be considered a collective asset that evolves with progress, not just a rigid regulatory tool.

Current research on language use among adolescents on social media ignores the impact of digital language practices on the preservation of formal Indonesian and regional languages. Although studies acknowledge the emergence of slang, code-switching, and the use of emojis, there has been insufficient research on how these practices affect national linguistic identity in the context of globalization. Furthermore, there has been no comprehensive research on the role of non-standard language forms, such as the combination of local languages, English, and Indonesian in code-switching, in shaping social identity. To understand how language development driven by young people through social media correlates with the preservation of local languages, further research is needed to explore the impact of these trends on language creativity and language policy.

II. LITERATURE REVIEW

Sociolinguistic studies examine language and society, focusing on how social variables such as age, gender, social class, and media influence language use. Language is not only a means of communication, but also serves as a representation of power, social identity, and culture, especially among the younger generation. This generation is a very active agent of language change, according to the style that shows their identity and according to the new language. Language use on social media confirms their social, cultural, and lifestyle affiliations.

As social media use continues to increase among teenagers, the phenomenon of language change on social media has become prominent among the younger generation. The emergence of slang, informal expressions, and code-mixing and code-switching, which are very popular on platforms such as Instagram, Facebook, and YouTube, is noteworthy. Social linguistics is useful for understanding how these social components influence language use in digital communication, especially how social groups on social media create their own language. The generalization of abbreviations and new terms when interacting on the internet (Atuzzahro et al., 2025). Language use on social media serves as a flexible and rapidly adapting channel of expression. This shows how informal language evolves more quickly in line with the changes of the current era.

Meanwhile, code-switching and code-mixing have become hallmarks of digital communication. Young people often switch between Indonesian, English, or even regional languages in a single sentence or conversation. This phenomenon occurs not only in spoken conversation, but also in written text on digital platforms, better known as code-mixing. This reflects the multilingual character of the Indonesian generation, who are fluent in both national and international languages, such as English. From a sociolinguistic perspective, this type of social identity has emerged as a result of globalization and advances in digital technology (Fadlilah et al., 2023).

However, the use of this new language also raises concerns about interest in Indonesian and regional languages. Standard Indonesian is increasingly marginalized in everyday communication due to the dominance of foreign languages on social media. Many see progress in linguistic creativity, but there are also those who worry that the younger generation will lose their regional and national linguistic identity. According to him, the use of regional languages in the media clearly shows the difference between language globalization and the preservation of regional

culture. This phenomenon reminds us of the danger of erosion of regional languages that have great cultural and historical value (Saddhono & Rohmadi, 2014).

Language change can reflect social change and group identity, not only in formal language. Language becomes very indicative of social identity in situations like this. It is not only a means of communication, but also a sign of unity and connection with a particular social group. Language is not only used by the younger generation, but by all groups with their respective social statuses. Language becomes a very important tool for distinguishing oneself from others in the digital world, where identity is often formed and displayed (Salman & Akhiroh, 2025; Silaban et al., 2025). For example, the phenomenon of slang usage on social media in Indonesia shows that language is not only a tool for communication but also a tool to show who they are and which group they represent (Zamhari et al., 2025). However, the components are constantly changing due to language changes on social media, but they also create and change the language to meet social and cultural needs. As a result, important terms in communication are created that may not exist in standard language. This shows that language on social media is greatly influenced by technological and cultural changes.

Sociolinguistics provides a better understanding of how language functions within a broader social context. It highlights the differences in how young people interpret words when they use them, as language serves as a representation of their evolving social identities. Furthermore, this shows how language functions as a tool to demonstrate power, create solidarity, and connect digital groups. Thus, understanding these dynamics is important for researchers, teachers, and policymakers. This will foster linguistic creativity without neglecting the preservation of language and local cultural identity (Saragih et al., 2025).

There are several gaps in existing research on language change among young Indonesians on social media. Although research has explored the use of slang, code-switching, and code-mixing, limited focus has been given to the impact of social media on the preservation of standard Indonesian and regional languages, especially among young people exposed to globalization. Research is needed to examine how social media influences language use and local cultural identity, especially on platforms such as Instagram, Facebook, and YouTube. In addition, the dominance of foreign languages, especially English, over Indonesian and regional languages on social media has not been widely studied. Research is needed on how to balance linguistic creativity with the preservation of regional languages in the digital age. This research can fill the gap in understanding the long-term impact of social media on cultural and linguistic identity in Indonesia.

III. METHODOLOGY

The qualitative descriptive method is the main approach used in this study. This method was chosen because it allows for obtaining in-depth and comprehensive information from the latest literature on the linguistic identity of the younger generation and language transformation in social media. This method is suitable for exploring the linguistic phenomena of Indonesia's younger generation on social media. This approach is used to understand the relationship between language and society, particularly in the context of language knowledge among the younger generation on social media (Muhid & Abdul, 2021).

Data collection was conducted through a literature review on social media platforms Facebook, Instagram, and YouTube that contained words or sentences with the following elements: use of slang and acronyms, code-mixing, use of emojis and punctuation, informal language style, use of hashtags, use of regional languages on social media, use of loanwords, use of words to express joy, use of sentences without subjects, and use of words to express courage. Various relevant library sources were used for data collection, including recent research on language change, particularly in the context of social media and language use by the younger generation (Hutauruk et al., 2025; Manurung et al., 2025). The literature search was conducted through digital platforms such as Google Scholar, DOAJ, and leading academic journals accredited in the fields of linguistics and sociolinguistics.

The data was analyzed descriptively and analytically, by collecting data on language transformation patterns occurring on social media and how this affects the linguistic identity of the younger generation. The focus of this study is how language is used to communicate social affiliation and self-identity in digital communication spaces. This analysis also examines how this can affect the preservation of standard Indonesian and regional languages, which are increasingly marginalized by the use of foreign languages and slang on social media (Putri & Sofyaningrum, 2024).

The analysis process involves triangulation of sources. This process compares relevant literature with other data to ensure that the analysis results are objective and unbiased, thereby ensuring the validity and reliability of the data (Sundari et al., 2025). The initial process is very important for understanding language transformation in social media and how it affects the linguistic identity of Indonesia's younger generation (Sembiring et al., 2025). This study aims to provide deeper insights into the language changes occurring on social media and how these changes affect the formation and negotiation of linguistic identity among the younger generation. Therefore, sociolinguistic research in this digital era can help policymakers develop plans to preserve the national language.

IV. RESULTS AND DISCUSSION

A. Language Change in Social Media

Young Indonesians use social media as their primary means of interacting with one another. The language used adapts very quickly during the communication process. English, in particular, is the most noticeable transformation among the younger generation. Indonesian, regional languages, and foreign languages. This is known as code-mixing. For example, sentences often appear in Facebook, Instagram, or YouTube posts:

"*Ide pose selfie biar photogenic.*"

'Selfie pose ideas to look photogenic'.

"*Gua pikir selama ini netflix and chill itu artinya nonton netflix sambil santuyyyy.*"

'I thought that "Netflix and chill" meant watching Netflix while relaxing'.

The examples in the sentence still partly use Indonesian, but words such as *selfie*, *Netflix*, and *chill* come from English and are now used daily. This shows how foreign languages, especially English, have entered and are used naturally in everyday conversation. The emergence of new slang terms among the younger generation has also brought about significant changes in language use on social media. Children have started using words such as *mantul*, which means 'really cool', *baper*, which means 'emotional,' and *galau*, which means 'restless'.

The linguistic variations used by Indonesia's younger generation on social media can be classified into ten forms, namely: the use of slang and acronyms, code mixing, the use of emojis and punctuation marks, informal language style, the use of hashtags, the use of regional languages on social media, the use of loanwords, the use of words to show affiliation, the use of sentences without subjects (ellipsis), and the use of words to show empowerment.

The following table presents findings on language practices in social media and their impact on linguistic identity.

TABLE 1
LANGUAGE PRACTICES ON SOCIAL MEDIA AND THEIR IMPACT ON IDENTITY

No.	Language Practice Category	Core Description	Sample Data	Sociolinguistic Functions	Impact of Identity	Risk Challenges
1	Use of Slang and Acronyms	slang lexicon, short emotional form	" <i>baper</i> = <i>bawa perasaan</i> " 'bring feelings' and " <i>cieee</i> " 'oh, my'	closeness index, solidarity	strengthening the youth in-group	blur the boundaries between standard and non-standard
2	Mix Code	Indonesia + English or regions in one utterance	" <i>ide hangout bareng bestie</i> " 'the idea of hanging out with friends'	global orientation, audience design	Cosmopolitan or multilingual identity	Marginalized local equivalents
3	Emojis and Punctuation	emojis, ellipses, repeated letters or symbols	" <i>alhamdulillah..</i> 🤗" 'praise be to God' and "🤗🤗"	visual prosody, mitigation of criticism	cohesion and emotional affiliation	cross-group ambiguity
4	Informal Style	abbreviations and morphological adaptations	" <i>pdhl udh nunggu loh??</i> ?" 'even though I've been waiting'	speed and familiarity	registered media style	social non-standard normalization
5	Hashtag Usage	tag series for visibility	"# <i>viral</i> # <i>pemula</i> # <i>interaksi</i> " # <i>viral</i> # <i>beginner</i> # <i>interaction</i>	visibility hacking, metadiscursive	creator or community identity	spam tag, performativity
6	Regional Language	local lexicon + glossary	" <i>ngabuburit</i> " 'waiting for the fast to break' and " <i>sekarepmu</i> " 'as you wish'	authenticity, translanguaging	ethnolinguistic pride	" <i>pemucatan</i> " 'bleaching' cultural meaning
7	Foreign Absorption	ensiklopedia urban-global	" <i>mal, proyek, restoran</i> " 'mall, project, restaurant'	modern or urban image index	professional identity or city	local equivalents eroded
8	Affective Expression	expression of affection or admiration	" <i>suka banget...</i> " 'really like' and " <i>...bikin gemes</i> " 'adorable'	phatic communion or stance	sense of togetherness	over-familiarity public
9	Ellipsis (Without Subject)	removal of clear elements from the context	" <i>yang baru mana</i> " 'which one is new' and " <i>di-prenk kah</i> " 'pranked'	language economics and particle prosody	common basic assumptions	ambiguous when the context broadens
10	Empowerment	inclusive or religious imperative	" <i>Jangan takut...</i> " 'don't be afraid' and " <i>Ayo kita masuk</i> " 'Let's go in'	collective or performative call	optimism and cohesion of action	risk of "toxic positivity"

(a) Use of Slang and Acronyms

"prank pake baju pacarku biking baper cewe kampung."

'Pranking by wearing my girlfriend's clothes made the village girl *baper*'.

"Keisya yang dielus kepalanya sama paul, satu Indonesia baper berjamaah."

'Keisya, whose head was stroked by Paul, one Indonesian *baper* together'.

"cie, baper cieee sampe bengong begitu."

'oh my, you're so *baper*, you're lost in thought like that'.

The current linguistic phenomenon that is popular among the younger generation is the word *baper*, which comes from the acronym '*bawa perasaan*' (Apyunita & Asdah, 2025). It shows commitment to a particular social group in slang used in sociolinguistics. Emotional slang on social media is stronger among its users. It is attractive because it is easy to use and can convey meaning that is quickly understood. From a sociolinguistic perspective, "*baper*" describes the younger generation in a more intimate and open way, making it a shared identity within their community (Saragih et al., 2025). However, this slang also does not indicate social boundaries or distinguish between people.

(b) Code Mixing

"Ide foto di pantai bareng bestie."

'Photo idea at the beach with your *bestie*'.

"Ide pose foto kece saat hangout."

'Cool photo pose ideas for *hanging out*'.

"Ide pose foto healing ke alam."

'Ideas for *healing* photo poses in nature'.

Code-mixing is the use of English words such as *bestie*, *hanging out*, and *healing* in Indonesian sentences (Sari et al., 2024; Fitri et al., 2023). Based on sociolinguistics, code-mixing is a form of language interaction that, according to the younger generation, is particularly prevalent on social media. This demonstrates the globalization of language. This change is particularly evident in English, especially in terms of lifestyle and trends, as a symbol of social status (Puspita & Setiana, 2025). On the other hand, code-mixing in language identity analysis allows people to demonstrate their understanding of global trends.

(c) Use of Emojis and Punctuation

"Alhamdulillah.. masi di kelilingi orang baik. 😊"

'Praise be to God, I am still surrounded by good people 😊'.

"Undangan suda mulai menumpuk.. entar jalan jalan 😊"

'The invitations are starting to pile up. Let's go for a walk 😊'.

"Aku potong rambut sendiri... 🪄🪄🪄."

'I cut my own hair 🪄🪄🪄'.

Nonverbal expressions, emojis, have become part of today's digital communication (Ramdani et al., 2025). The use of emojis 😊 and 🪄🪄🪄 enhances the message being conveyed and adds a dimension that cannot be achieved with words alone. This shows in sociolinguistics how nonverbal communication is an important component of language on social media. Social media users employ emojis because they consider them a more direct and effective way to convey their psychosocial feelings (Sembiring et al., 2025). Conversely, social identity analysis finds that the widespread use of emojis on social media also indicates a connection to specific digital communities, resulting in broader identities and increased usage.

(d) Informal Style

"Pdhl aku udh nunggu loh...knp ini di post lagi??"

'Even though I've been waiting, why is this being uploaded again?'

"Pantes dejavu, ini eps udh nonton di ulang lagi"

'No wonder I had déjàvu, I've already watched this episode again'.

"Uplod tiap hari juga gapapaa, seneng bgt liat mrkaaaaaa"

'It's okay to upload every day, I'm so happy to see them'.

The three statements above show a style shift towards a casual register, indexed through abbreviations (*pdhl*, *udh*, *knp*, and *bgt*), English loanwords (*post*, *upload*, *eps*), and grapheme lengthening (*loh* and *mrkaaaaaa*), as well as double question marks. Within the framework of indexicality, these forms function as indices of familiarity while also marking the identity of social media users. The variation in *di post* also shows spontaneous morphological adaptation to loanwords, evidence of efficient translanguaging that fosters solidarity. Through the lens of audience design, this stylistic choice is clearly oriented toward a peer audience, so that speakers adjust their register to align with community norms and the fast-paced expectations of the platform. This process supports the enregisterment of social media language features, collectively known as a style that is appropriate in a particular space, thereby strengthening a sense of solidarity (Skujins, 2017).

From a social function perspective, these forms of language demonstrate positive politeness strategies and facework. Complaints such as "why post it again?" are conveyed in a casual tone, while comments such as "déjà vu" are not offensive, and statements such as "it's okay to upload every day" show support. In digital communication, punctuation, letter elongation, and ellipsis replace tone of voice or facial expressions. From the perspective of *Communities of*

Practice, habits reveal how abbreviations, code-mixing, and exaggerated styles become characteristic of social media user communities to maintain familiarity and strengthen a sense of togetherness (Tafsir, 2023).

(e) *Hashtag Usage*

"#sorotan #gastag #jangkauanpublik #viral #fbpro #semuaorang #pemula #pengikut #gastag #semuanya selamat petang #kontenkreitor semoga kita cepat #dimonetisasi dgn #gastag yg baik #edukasi ok Mari sama2 #interaksi sesama #pemula."

'#highlights #gastag #publicreach #viral #fbpro #everyone #beginners #followers #gastag #everything good afternoon #contentcreators hopefully we will quickly #monetize with #gastag that is good #education ok Let's #interact with each other #beginners.'

The series of hashtags above show a style shift to a marketing-community register that is characteristic of the platform ecosystem (Susilawati, 2024). Indexically, these hashtags mark: the economic orientation of the platform (wanting to appear or be seen); membership position (who we are); interaction goals (what we do together). The use of #gastag serves as a rhythmic enhancer as well as a visibility hacking strategy with the aim of increasing the chances of penetrating the algorithm through the duplication of community labels. The mix of informal orthography: the numeronym "sama2," the abbreviation "dgn," and the greeting "selamat petang" with its Malay flavor, shows translanguaging across Indonesian-Malay varieties that affiliates speakers at the regional level, while giving a non-hierarchical and friendly impression (Mohamad, 2020).

From the perspective of audience design and communities of practice, the series of hashtags plus the invitation "*let's do it together*" positions the speaker as a peer who invites participation, not as an authority—thus fitting in with the engagement-based attention economy (Skujins, 2017). A number of hashtags are metadiscursive: they not only index topics, but also instruct how to behave on the platform, namely learning, interacting with each other, and pursuing monetization. The phatic function is strong, so that hashtags become a "circuit greeting" to ensure that the channel remains alive; while the performative function of identity makes hashtag language a symbolic uniform for novice creators. Thus, the hashtag information in this data is not just a casual style. It is a social-algorithmic device for expressing membership, negotiating aspirations for growth and monetization, and mobilizing collective engagement in the logic of trending (Bajuri et al., 2024).

(f) *Use of Regional Languages on Social Media*

"Bismillah ngabuburit sambil banyak tanya 🤔"

'In the name of Allah, (ngabuburit = waiting for the time to break the fast) while continuing to wonder'.

"Ora usah khawatir ora dadi opo-opo...."

'Don't worry, nothing will happen'.

"... ya wess sekarepmu ae kajej ahh..."

'... Yes, whatever, do as you please, respectfully'.

Regional language lexicon such as *ngabuburit*, *ora dadi opo-opo*, *sekarepmu*, and *kajej* are not merely a choice of words, but also markers of identity and attitude of speakers in the digital space. This phenomenon often occurs in captions, comments, and memes on various social media platforms, forming a distinctive register that reflects local authenticity and cultural solidarity. Non-standard orthography is always used as a signal of regionalism, while short glosses—for example, "*ngabuburit*: waiting for the fast to break"—expand access to meaning across online communities. Thus, regional languages not only have a communicative function but also serve as symbolic capital and a representation of cultural identity in the current digital era (Pawestri, 2019).

With its wide reach and algorithmic boost, social media has created a new space for translanguaging practices. There is a mixture of regional languages, Indonesian, and foreign languages, making communication more relevant among the younger generation. This practice strengthens cultural bonds among local users while bridging audiences across regions. However, there is a risk of "meaning dilution" when regional lexicon is treated as a style without a fundamental understanding of the culture. Therefore, digital literacy that involves awareness of cultural context, gloss, and language ethics is necessary. Collaboration between creators, local communities, and social media platforms can strengthen the function of digital media as a means of adaptive and sustainable preservation of regional languages (Lestari et al., 2025).

(g) *Use of Loanwords*

"Beginilah akibat kalau punya kendaraan tapi tidak tahu aturan dan adab parkir. Posisi persis di depan Mall Panakkukang."

'This is what happens when you own a vehicle but don't know the rules and etiquette of parking. Positioned right in front of Panakkukang Mall'.

"Dibutuhkan PELAKSANA PROYEK untuk pembangunan rumah tinggal. Proyek-proyek berlokasi di Jakarta"

'PROJECT MANAGERS are needed for residential construction projects. Projects are located in Jakarta'.

"Melihat reaksi Rifin Dendeng & Markonah makan di restoran"

'Seeing Rifin Dendeng & Markonah's reaction to eating at a restaurant.'

The use of loanwords with Indonesian language usage patterns such as *mall*, *project*, and *restaurant* is not merely a lexical choice but functions as an indexical marker positioning the speaker in the image of a global identity: cosmopolitan, modern, and connected across cities and countries. In communicative practice, these words have been registered as part of a prestigious urban style laden with symbolic value. The word *mal* signifies middle-class

consumption and leisure discourse; *proyek* refers to the spirit of professionalism, work mobility, and socio-economic networks; while *restoran* represents a culinary lifestyle that transcends geographical and cultural boundaries. Variations in form, such as *mal* versus *mall* or *restoran* versus *resto*, as well as the use of hashtags and emojis, show how speakers negotiate their self-image—whether formal, casual, or trend-aware. These choices are performative, as they relate to efforts at impression management in front of an audience and involvement in a digital economy that highly values global symbols (Elisa et al., 2024).

The use of foreign vocabulary in the digital space brings with it certain values—efficiency, innovation, lifestyle, and modernity—which are reinforced through circulation in *captions*, reviews, and *memes*. Through this process, a new understanding of what is considered modern and valuable in digital society is formed. Social media algorithms further reinforce this trend: the key is to be global in nature, increasing searchability, expanding reach, and normalizing usage to the point of becoming a stylistic standard. This phenomenon demonstrates the dialectical character of lexical globalization—on the one hand, it enriches the linguistic repertoire and supports the practice of translanguaging (Indonesian-English/foreign-borrowed/regional), but on the other hand, it has the potential to marginalize local equivalents and weaken the function of national and regional languages. This situation requires critical and adaptive language literacy, which is sensitive to context, dominance, and audience, so that users can assess when global terms are effective in building image and access, and when local equivalents are more appropriate for clarity and inclusivity of linguistic identity (Luwiti, 2025). With this framework, the use of *mall*, *project*, and *restaurant* on social media is not merely an imitation of globality, but a conscious strategy to negotiate meaning, value, and social position in the dynamic landscape of digital communication.

(h) *Use of Words to Show Affection (Affiliation)*

"Ohh bgini toh aslinya 😊"

'Oh. So that's how it really is 😊'.

"Suka bngt sm BBK,,, pliss rajin2 bikin vlog yak"

'I really like BBK, please keep making vlogs, okay'

"Unyu-unyu gimana gitu 17 tuh, bikin gemes."

'It's so cute at the age of 17, it makes you feel affectionate'.

Expressions such as 😊, *wkwkwk*, and popular phrases such as *oh begitu toh aslinya*, "*suka banget...*" atau "...*bikin gemes*" play an important role as effective tools that restore nonverbal cues in digital communication. Sociolinguistics views these expressions as signs of the speaker's stance or attitude, which index closeness, surprise, admiration, or social intimacy in the informal online realm. Graphemic features such as ellipsis, repeated letters such as the word "*bangett*," selective capitalization, and emoji choices act as visual prosody that give rhythm and mark the intensity of the message in the text. Variations of laughter such as *wkwkwk* and *xixixi* or *lol* indicate patterns of community membership (in-group markers), which strengthen group identity while mitigating potential face-threatening acts through humor and familiarity. Thus, emojis and digital forms of laughter are not only decorative, but also function as social symbols that refine emotional expression and strengthen linguistic solidarity on social media (Ramdani et al., 2025; Sembiring et al., 2025).

In the realm of social relations, emojis and digital emotional expressions function as phatic communion, a form of communication that maintains social relationships and the continuity of interaction, rather than merely conveying information. Through the emotional connection in speech (affective alignment), the use of social media is able to negotiate closeness and solidarity in the context of context collapse, where the boundaries between public and private audiences become blurred. Meanwhile, at the online community level, this pattern develops into a style norm (register): when laughter is considered friendly, when ellipsis means hesitation, and when intensifiers such as the words *banget* or *parah sih* reinforce social participation. Thus, emojis act as pragmatic capital that becomes social cohesion and reinforces group identity in the rapidly changing digital world (Ilham et al., 2025; Nisa, 2025).

(i) *Use of Sentences Without Subjects (Ellipsis)*

"yang baru mana"

'which one is new'

"di prank kah"

'was it a prank'

"Ohh bgini toh aslinya 😊"

'Oh, so that's how it originally is 😊'.

The omission of subjects in utterances such as *yang baru mana* (which one is new), *di-prank kah* (was it a prank), or *ohh begini toh aslinya* (oh, so that's how it really is) 😊 is a very common form of ellipsis strategy in digital conversations. Typologically and pragmatically, Indonesian is a *topic-prominent* language, which allows the subject to be omitted when the topic is already salient from the context of the conversation. In the realm of high-speed social media, ellipsis functions as an economy of language: it saves typing, speeds up responses, and adapts to linear conversation formats such as *threads*, comments, or direct messages that have a strong co-textual context. Nonverbal unities such as interjection emojis (*ohh*), modal particles (*kah*, *toh*), and creative orthographic variations function as *visual prosody* that fills grammatical information gaps: indexing attitudes (surprise, playfulness, or empathy), marking

speech acts (questioning vs. stating), and regulating social distance without the need to explicitly mention the subject. This phenomenon shows that the strategy of ellipsis in digital space is not merely a simplification of syntax, but rather a communicative adaptation tailored to the speed, context, and style of online interaction (Iswatiningsih et al., 2024; Sembiring et al., 2025).

From a sociolinguistic perspective, ellipsis can also be interpreted as a stylistic choice that signifies familiarity and solidarity within the online community. Sentences without subjects indicate the existence of an imagined audience that shares common ground, so that the omission of sentence elements does not reduce the comprehensibility of meaning. In the context of digital conversation, this form reinforces the phatic function, which is to maintain continuity of interaction and a sense of closeness among participants. However, language economy carries the risk of ambiguity when the context of the conversation shifts or the audience broadens. Therefore, users often add clarifying elements such as repetition of objects (the new black), topic markers (#tags), or expressive emojis to maintain relevance and meaning. Thus, ellipsis in digital communication can be understood not only as a form of economy, but also as a pragmatic strategy to balance speed, efficiency, and social familiarity in online communities (Nababan et al., 2024).

(j) *Use of Words to Show Courage (Empowerment)*

"*Jangan Takut: Sebab Allah Menyertai Engkau*"

'Do Not Be Afraid: For God Is With You'

"Ketakutan yang berlebihan kadang membuat kita tidak sadar dalam bertindak. Tingkah laku yang tidak disadari karena ketakutan itu bisa menimbulkan penilaian buruk dari orang lain."

'Excessive fear sometimes makes us unaware of our behavior. Unconscious behavior caused by fear can lead to negative judgments from others'.

"*Temani Admin Masuk Goa Putri Asih Tuban. Yang takut hantu jangan lihat video ini.. Ayo kita masuk....*"

'Join the Admin as we enter Putri Asih Cave in Tuban. If you're afraid of ghosts, don't watch this video. Let's go inside'.

Expressions such as *don't be afraid: because God is with you, excessive fear*, or *let's go in* are forms of speech that function as markers of productive empowerment on social media. Pragmatically, imperative forms such as *don't be afraid* or *let's go...* position the speaker as a motivator who provides moral encouragement and frames risk as a shared challenge. In addition, religious intertextuality, as in the quote "God is with you," gives moral and spiritual legitimacy to the utterance, reinforcing its persuasive effect. From a stance perspective, this lexical choice marks two dimensions of attitude at once: deontic, which is assigning oneself or a group to act courageously; and affective, which is calming collective anxiety. Multimodal features such as enthusiastic emojis, selective capital letters (*JANGAN TAKUT*), or ellipsis signs reinforce the process of courage by adding emotional intensity to the text. Performatively, the utterance *don't be afraid* not only reports an emotional condition but also creates a new social condition, namely instilling confidence and activating motivation in the imagined audience (Kareem, 2023; Hashmi et al., 2022).

From a sociolinguistic perspective, this discourse of courage plays an important role in shaping social identity and group cohesion in the digital space. Inclusive imperatives such as the word *we* and participatory narratives such as *accompany the admin* foster solidarity and *affective alignment*, which is the emotional alignment between users that reinforces a sense of togetherness. Social media algorithms also reinforce speech with actionable intonation, encouraging the rapid spread of positive vocabulary such as "optimism," "never give up," and "dare to try." However, this phenomenon is also dialectical: excessive encouragement of courage can lead to "toxic positivity," which is an overemphasis on enthusiasm that ignores the reality of risk and emotional vulnerability. Therefore, reflective courage literacy is important, which involves the use of language that motivates without erasing empathy or cultural context. In this framework, statements like *don't be afraid* can become an ethic of courage that is realistic, empathetic, and empowering for the digital community (Arndt, 2022; Ross & Bhatia, 2025).

B. Impact on Linguistic Identity

The use of language on social media greatly influences the formation of linguistic identity among the younger generation, where language is used as a tool to express oneself and reinforce membership in certain social groups (Rianto & Juliarta, 2024). Code-mixing, the use of slang, and informal language on social media are characteristics that reinforce the social identity of young people (Wijaya, 2022). Posting statuses with slang terms and social media language is a form of adaptation to the times, as well as a way of affirming involvement in online communities (Setyowati & Yanottama, 2023). The process of linguistic transformation that occurs shows the flexibility of language use in the digital age and demonstrates how social media has become a space for negotiating personal and group identities (Sulaiman et al., 2025).

Social media has provided opportunities for multilingual identities. Young people, who often use *multiple* languages in their daily lives, are beginning to show flexible linguistic identities. This ability allows them to use different languages according to the platform and audience, enabling them to adapt to various social environments, both formal and informal. Their language use is multidimensional, as seen in their ability to switch from formal Indonesian in professional settings to informal slang and foreign terms (especially from English) in online communication (Dewaele & Wei, 2014).

It appears that linguistic adaptability is becoming increasingly relevant as part of their digital identity, along with a growing awareness of the multicultural environment around them (Arafah & Hasyim, 2019). In addition to adaptability,

code-switching and code-mixing have become a major focus in how young people shape their identities (Kay et al., 2022). Switching between Indonesian, English, and sometimes regional languages in a single post or conversation, for example, demonstrates a person's ability to easily mix languages depending on the context, audience, and purpose (Rahman & Muttaqin, 2024). This linguistic landscape is further enriched by the development of new linguistic forms such as internet memes, abbreviations, and emojis, providing new modes of social expression unique to the digital generation (Duek & Nilsberth, 2024). In digital communication, the use of emojis is always present because emojis are considered one of the new linguistic features that have emerged in digital communication. Emojis serve as expressive components that complement text and help convey specific meanings, emotions, or nuances in short messages. Emojis have become part of a growing linguistic phenomenon as young people use technology and social media.

The dominance of foreign languages, especially English, on social media has raised questions about the erosion of the Indonesian language, particularly the use of local dialects and regional languages (Nurpratiwiningsih & Maknun, 2020). Indonesian, which is a symbol of national identity, is at risk of being marginalized as English becomes increasingly integrated into digital interactions. The use of English words or Anglicized slang has become so common that it threatens linguistic diversity in Indonesia. Many young people, especially in big cities, often choose English words and phrases without realizing the impact on their mother tongue or country. As young people choose the convenience and global appeal of English, this can lead to a gradual decline in regional languages. Especially among the younger generation who grow up with little exposure to regional languages, the loss of local languages can lead to the loss of ethnolinguistic identity and cultural nuances (Siregar et al., 2025).

These languages have great cultural and historical value, so their gradual extinction could threaten Indonesia's rich linguistic diversity. For example, expressions of cultural practices, customs, and community values are highly dependent on regional languages. Therefore, the general dominance of English and the lack of attention to the use of regional languages on social media raise important issues regarding the future of linguistic diversity and the need for policies to preserve and develop regional languages in the digital space (Cahyani et al., 2021).

Finally, the use of language on social media has made it a useful tool for the younger generation in shaping their linguistic identity. This poses a major challenge, especially in terms of preserving Indonesian and regional languages, even though it has encouraged flexibility, creativity, and the development of new forms of digital communication. The growing influence of English and the dominance of digital communication forms demand a balanced approach that guarantees linguistic innovation and ensures that the richness of regional language heritage is not lost in the process. The implementation of language policies that support inclusive multilingualism and guarantee the preservation of linguistic diversity in the digital age is essential as the digital environment changes.

V. CONCLUSIONS AND IMPLICATIONS

Language use among Indonesia's younger generation is changing rapidly and dynamically, especially on social media. In everyday interactions, forms such as slang, code-mixing, code-switching, and the use of foreign vocabulary—especially English—are becoming increasingly common. These linguistic changes are not merely a matter of style, but are an important tool in identity formation, allowing young people to express their goals, identify with certain social groups, and reflect urban or global mindsets. As a hallmark of their digital identity, Indonesia's young generation now shows great flexibility in switching between Indonesian, English, and regional languages according to the situation. Standard Indonesian and regional languages face marginalization due to the prevalence of informal urban Indonesian and the dominance of English in the digital space. Linguistic diversity and cultural nuances may be gradually eroded by this trend. As a result, social media has become a new arena for negotiating linguistic ideology and seeking a balance between the demands of globalization and the need to preserve the national language.

These findings have important implications. The results emphasize how important it is for the government and educational institutions to promote literacy and language awareness. The aim is to encourage the simultaneous use of standard Indonesian and regional languages, as well as sustainable linguistic innovation, rather than to hinder creativity. In addition to initiatives to preserve and revive regional languages by integrating them into digital literacy programs for the younger generation, inclusive language policies are needed to encourage the adaptive use of language in the online context. Furthermore, proactive measures to protect Indonesia's linguistic and cultural diversity in the digital age are urgently needed, given the potential loss of cultural identity due to language loss. To understand these ongoing changes and their broader impact on the identity of the younger generation, this study also encourages more empirical sociolinguistic research, including corpus studies and digital discourse analysis. Thus, social media encourages language innovation and identity negotiation among Indonesian youth, but also poses challenges for language preservation and cultural heritage, requiring thoughtful and balanced interventions to support innovation while preserving Indonesia's rich linguistic diversity.

ACKNOWLEDGEMENTS

My gratitude is extended to the *Lembaga Pengelola Dana Pendidikan (LPDP)* for funding this research and for supporting the second authors, continued in the Doctoral Program in Linguistics at Hasanuddin University.

REFERENCES

- [1] Apyunita, D., & Asdah, A. N. (2025). Representation of Slang Among Generation Z on Instagram Social Media. *DEIKTIS: Jurnal Pendidikan Bahasa Dan Sastra*, 5(2), 1080–1086. <https://doi.org/10.53769/deiktis.v5i2.1608>
- [2] Arafah, B., & Hasyim, M. (2019). The Language of Emoji in Social Media. *KnE Social Sciences*, 2019, 494–504. <https://doi.org/10.18502/kss.v3i19.4880>
- [3] Arévalo-Salinas, A. I., García-López, M., Blanco-Morett, Á., & Simancas-González, E. (2025). TikTok as a communication and information tool for Spanish local councils. *Technology in Society*, 82(March). <https://doi.org/10.1016/j.techsoc.2025.102885>
- [4] Arndt, R. (2022). Vocabulary in digital science resources for middle school learners. *Applied Corpus Linguistics*, 2(3), 1–9. <https://doi.org/10.1016/j.acorp.2022.100023>
- [5] Asdah, A. N., & Safitri, N. A. S. (2025). Code-Mixing Practices in Digital Interactions of Indonesian Language and Literature Education Students on Social Media: A Sociolinguistic Study. *DEIKTIS: Jurnal Pendidikan Bahasa Dan Sastra*, 5(2), 906–918. <https://doi.org/10.53769/deiktis.v5i2.1590>
- [6] Atuzzahro, A. A., Handatika, D. T., & Puspita, A. R. (2025). The Use of Slang in Digital Communication on TikTok Content by Nathalie Tresya. *BLAZE: Jurnal Bahasa Dan Sastra Dalam Pendidikan Linguistik Dan Pengembangan*, 3(2), 169–178. <https://doi.org/10.59841/blaze.v3i2.2737>
- [7] Bajuri, M. Z. J., Rahman, F., & Ilma, A. A. (2024). Language development in social media: from slang to popular abbreviations. *Pujangga: Jurnal Bahasa dan Sastra*, 10, 150–166.
- [8] Biagi, I. W. K. Di, & Suryadmaja, G. (2025). Trends in Slang Use Among Teenagers on the TikTok Platform: Analysis of Meaning and Distribution. *Journal of Linguistics and Language Teaching*, 1(1), 22–29. <https://doi.org/10.71094/jollt.v1i1.47>
- [9] Cahyani, M. I., Agan, S., & Lailiyah, N. (2021). The Influence of English on the Use of Indonesian in Teen Communication on Twitter and Instagram. *Prosiding SEMDIKJAR*, 4, 450–462.
- [10] Dewaele, J.-M., & Wei, L. (2014). Attitudes towards code-switching among adult mono- and multilingual language users. *Journal of Multilingual and Multicultural Development*, 35(3), 235–251. <https://doi.org/10.1080/01434632.2013.859687>
- [11] Duek, S., & Nilsberth, M. (2024). Languages and identity constructions in multilingual students' digital literacy practices. *Education Inquiry*, 15(3), 295–311. <https://doi.org/10.1080/20004508.2022.2109561>
- [12] Elisa, L., Rahman, B., & Sinaga, R. M. (2024). The Transformation of Indonesian Identity among IAI Tulangbawang Students in the Use of English as a Global Language. *Bulletin of Community Engagement*, 4(3), 1–9.
- [13] Fadlilah, A., Utari, R., & Iskandar, D. (2023). Language and Group Identity in Political Figures' Statements: A Sociolinguistic Study. *Sasdaya: Gadjah Mada Journal of Humanities*, 7(2), 172–189. <https://doi.org/10.22146/sasdaya.10200>
- [14] Fitri, A. N., Firmanda, A., Ananda, D. T., Aulia, S., Setiawan, W., & Julianti, U. (2023). Mapping English-Indonesian Code-Mixing on Social Media: Its Impact on Digital Literary Language. *Jurnal Multidisiplin Saintek*, 45(1), 1–17.
- [15] Hashmi, U. M., Hashmi, H. A., Rashid, R. A. B., Shahzad, M., Almekhlafy, S. S. A., & Malik, M. Y. (2022). Former Muslims' Socio-Religious Discourse on Social Media: A Speech Acts Analysis. *3L: Language, Linguistics, Literature*, 28(4), 197–211. <https://doi.org/10.17576/3L-2022-2804-14>
- [16] Hutaaruk, D. D. R., Banjarnahor, E. T., Siahaan, W. A., & Surip, M. (2025). The Dynamics of Youth Language on Social Media: A Case Study of TikTok and Instagram. *Sosial Dan Humaniora*, 1(3), 457–467.
- [17] Ilham, M. A., Supriadi, R., & Farisi, M. Z. A. (2025). The Transformation of Indonesian Language in the Digital Context: Changes in Language Structure and Form. *Jurnal Bastra (Bahasa Dan Sastra)*, 10(1), 1–11.
- [18] Iswatiningsih, D., Fauzan, & Pangesti, F. (2021). Millennial youth expression through the use of slang on social media. *KEMBARA Journal of Scientific Language Literature and Teaching*, 7(2), 476–489. <https://doi.org/10.22219/kembara.v7i2.18301>
- [19] Iswatiningsih, D., Melati, I. K., & Zahidi, M. K. (2024). The Dynamics of Visual and Digital Language Among Generation Alpha in Everyday Communication on Social Media. *Paramasastra*, 11(2), 322–338. <https://doi.org/10.26740/paramasastra.v11n2.p322-331>
- [20] Kareem, R. H. (2023). A Pragma-Rhetorical Study of Motivation In Selected Quranic Verses. *Journal of Human Sciences*, 14(4), 5008-5027. <https://doi.org/10.33855/0905-030-004-001>
- [21] Kay, A. Y. A., Nitiasih, P. K., & Suarnajaya, I. W. (2022). the Analysis of the Uses of Code Switching and Code Mixing in Social Media Among Facebookers. *Jurnal Pendidikan Bahasa Inggris Indonesia*, 10(1), 1–14. <https://doi.org/10.23887/jpbi.v10i1.849>
- [22] Kittiveja, O., Srinon, U., & Srimuang, P. (2025). Analysis of Linguistic Strategies in Naming Restaurants and Street Food Stalls in Thailand: A Comparative Linguistic Landscape. *Forum for Linguistic Studies*, 7(6), 1167–1180. <https://doi.org/10.30564/fls.v7i6.9126>
- [23] Lestari, A. S. A., Anwar, W. P., Sibali, A., Misnawati, M., & Musa, N. A. (2025). Translanguaging as an Innovative Strategy in English Language Learning: Theory, Implementation, and Challenges. *Journal of Indonesian Scholars for Social Research*, 5(1 Special Issue), 1–13. [https://doi.org/10.59065/jissr.v5i1%20\(Special%20Issues\).181](https://doi.org/10.59065/jissr.v5i1%20(Special%20Issues).181)
- [24] Luwiti, S. R. (2025). The Impact of Social Media on Language Shift in Gorontalo Among the Younger Generation: Content Analysis of TikTok and Instagram. *SR Luwiti Reduplikasi: Jurnal Penelitian Pendidikan Bahasa Indonesia*, 5(1), 25–33.
- [25] Manurung, E. P. N., Manik, J. L. K., Sibuea, G. S., & Surip, M. (2025). The Dynamics of Indonesian Language in the Digital Age: Between Linguistic Development and Preservation Challenges. *Morfologi: Jurnal Ilmu Pendidikan, Bahasa, Sastra Dan Budaya*, 3(2), 248–257. <https://doi.org/10.61132/morfologi.v3i2.1551>
- [26] Mardikantoro, H. B., Siroj, M. B., Utami, E. S., & Kurniati, E. (2023). Investigating Indonesian language varieties in social media interactions: Implications to teaching practices. *Indonesian Journal of Applied Linguistics*, 13(2), 306–316. <https://doi.org/10.17509/ijal.v13i2.63069>
- [27] Mohamad, S. (2020). *Social meaning, indexicality and enregisterment of Manglish in youth WhatsApp* [Doctoral dissertation]. University of Liverpool, Inggris. Retrieved December 22, 2025, from

- https://livrepository.liverpool.ac.uk/3086489/1/201066019_Sept2019.pdf
- [28] Muhid, A., & Abdul, H. Y. (2021). *Sociolinguistics: approach and theory*. Jawa tengah: ganesha kreasi semesta.
 - [29] Nababan, W. R., Rahmadani, N., Tamba, W. O. V., & Hidayat Nst, T. K. (2024). Language Challenges in the Digital Age Regarding Language Errors in Social Media Communication. *Jurnal Bahasa Daerah Indonesia*, 1(3), 1–9. <https://doi.org/10.47134/jbdi.v1i3.2602>
 - [30] Natsir, N., Aliah, N., Zulkhaeriyah, Z., Amiruddin, A., & Esmianti, F. (2023). The Impact of Language Changes Caused By Technology and Social Media. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 7(1), 115–124. <https://doi.org/10.30743/ll.v7i1.7021>
 - [31] Nisa, Y. K. (2025). Semiotic Analysis of Emoji Use in Online Communication by Teenagers. *Jurnal Ilmiah Research Student*, 2(2), 852–859. <https://doi.org/10.61722/jirs.v2i2.5876>
 - [32] Nurpratiwiningsih, L., & Maknun, M. J. (2020). The Impact of Globalization on the Use of Indonesian Language by the Community. *Jurnal KONTEKSTUAL*, 1(2), 43–48.
 - [33] Pawestri, A. G. (2019). Building Banyumasan Cultural Identity through The Ngapak Dialect on Social Media. *Jurnal Pendidikan Bahasa Dan Sastra*, 19(2), 255–266. https://doi.org/10.17509/bs_jpbs.v19i2.24790
 - [34] Puspita, S. Della, & Setiana, L. N. (2025). Language Variation on Timothy Ronald's YouTube Channel "You Must Be Rich Now" (Sociolinguistic Study). *Jurnal Pendidikan Bahasa Indonesia*, 13(1), 60–70. <https://doi.org/10.30659/jpbi.13.1.67-76>
 - [35] Putri, I. T., & Sofyaningrum, R. (2024). The Meaning and Use of Slang by Gen Alpha in Online Communication in the Era of Society 5.0. *DEIKTIS: Jurnal Pendidikan Bahasa Dan Sastra*, 4(4), 871–886. <https://doi.org/10.53769/deiktis.v4i4.1180>
 - [36] Rahman, L. I., & Muttaqin, D. A. (2024). Instagram Captions by Berlian College Students: A Blend of Indonesian and English Code-Mixing. *Jurnal Ilmiah Insan Mulia*, 1(2), 63–71. <https://doi.org/10.59923/jiim.v1i2.298>
 - [37] Ramdani, D., Sholehah, N. F., & Nandang, A. (2025). Emojis as a Social Communication Strategy: A Sociolinguistic Analysis of Digital Texts. *Emotion*, 5(3), 2660–2669.
 - [38] Rianto, S., & Juliarta, I. M. (2024). The Impact of Social Media on Language Use Among Teenagers. *Focus Journal Language Review*, 2(2), 33–45. <https://doi.org/10.62795/fjlg.v2i2.297>
 - [39] Ross, A. S., & Bhatia, A. (2025). @Fesshole and the discourse of confession on X: A study of online sharing and community building. *Discourse, Context and Media*, 63(12), 1–8. <https://doi.org/10.1016/j.dcm.2024.100838>
 - [40] Saddhono, K., & Rohmadi, M. (2014). A sociolinguistics study on the use of the Javanese language in the learning process in primary schools in Surakarta, Central Java, Indonesia. *International Education Studies*, 7(6), 25–30. <https://doi.org/10.5539/ies.v7n6p25>
 - [41] Salman, A. D., & Akhiroh, N. S. (2025). The Use of "Jaksel" Language as Slang Among Gen Z in South Jakarta. *JlIP - Jurnal Ilmiah Ilmu Pendidikan*, 8(9), 10361–10368. <https://doi.org/10.54371/jiip.v8i9.9141>
 - [42] Saragih, E. L. L., Simorangkir, N. A., Artonang, C., & Marbun, O. D. (2025). The Use of Slang Among Millennial Teens on Twitter. *MUARA PENDIDIKAN: Jurnal Ilmiah Ilmu Pendidikan & Sosial Humaniora*, 1(3), 33–45. <https://doi.org/10.64365/muradik.v1i3.28>
 - [43] Sari, N. I., Asy'arie, B. F., Jamilah, Fitriyah, & Ridho, M. (2024). Internalization of Islamic Education Values in a Cultural Frame: A Systematic Literature Review in Wayang Art. *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan*, 22(2), 201–224. <https://doi.org/10.21154/cendekia.v22i2.9918>
 - [44] Sembiring, M. E. B., Hutabarat, S., Sinaga, D. A. C., & Wulan, E. S. P. (2025). Language Variation in Digital Communication: A Sociolinguistic Study on Social Media Platforms. *Jurnal Ilmu Pengetahuan Naratif*, 6(1), 85–95.
 - [45] Setyowati, R., & Yanottama, F. (2023). An Analysis of Slang on Twitter in Indonesia. *Proceeding of International Conference on Science, Health, And Technology*, 522–531. <https://doi.org/10.47701/icohetech.v4i1.3437>
 - [46] Silaban, N., Silaban, P., Simamora, F., Nadeak, O., & Sari, Y. (2025). Changes in Word Meaning in Teenage Language: A Semantic Review of Slang in Digital Media. *Jejak Digital: Jurnal Ilmiah Multidisiplin*, 1(4), 696–703. <https://doi.org/10.63822/0je4zq38>
 - [47] Siregar, A. M., Zai, E., Juana, M., Hutagaol, S. M. A., Siagian, Y. Y., & Tansliova, L. (2025). The Influence of Foreign Languages on the Existence of Indonesian: A Case Study on the Use of Slang and Mixed Languages. *Jurnal Akuntansi Hukum Dan Edukasi*, 2(1), 98–106. <https://doi.org/10.57235/jahe.v2i1.5793>
 - [48] Sitanggang, G. A., Sembiring, A. B., Nafisa, S., Syaughia, S., Syam, H. M., Saragih, C. A., & Anggreani, A. (2025). Analysis of Language Change and the Formation of New Variants in Social Media. *Jurnal Intelek Dan Cendekiawan Nusantara*, 2(4), 7343–7346.
 - [49] Skujins, C. (2017). *Indonesian/English Code-Switching On Social Media* [Master Thesis]. Flinders University, Australia. Retrieved December 23, 2025, from <https://flex.flinders.edu.au/file/ca9ee831-31e4-435e-aa543a34dbade3eb/1/C.SKUJINS%20Masters%20Thesis%20v21.pdf>
 - [50] Sulaiman, M., Santoso, A., & Susanto, G. (2025). Generation Z's prokem in social media: Language transformation and social identity of Makassar Adolescents. *Satwika: Kajian Ilmu Budaya Dan Perubahan Sosial*, 9(1), 17–28. <https://doi.org/10.22219/satwika.v9i1.37487>
 - [51] Sundari, Sutardi, & Mustofa. (2025). Tiktok Social Media and the Development of Communication Language in Elementary School Students: A Psycholinguistic and Sociolinguistic Study. *Jurnal Bahasa, Sastra, Pendidikan dan Humaniora*, 2(1), 69–79.
 - [52] Susilawati. (2024). English as a Marker of Social Identity Among Riau Youth. *Al-Amin: Jurnal Ilmu Pendidikan Dan Sosial Humaniora*, 2(9), 73–84.
 - [53] Tafsir, M. H. (2023). *Code-Switching Used By the English-Speaking Comedians* [Master Thesis]. Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia. Retrieved December 22, 2025, from <http://etheses.uin-malang.ac.id/51545/1/18320168.pdf>
 - [54] Tarihoran, N., Fachriyah, E., Tressyalina, & Sumirat, I. R. (2022). The Impact of Social Media on the Use of Code Mixing by Generation Z. *International Journal of Interactive Mobile Technologies*, 16(7), 54–69. <https://doi.org/10.3991/ijim.v16i07.27659>
 - [55] Wijaya, D. D. (2022). Code-Mixing on Instagram Paptions Posted by The Fashion Influencer "AMELIA ELLE." *Journal of*

English for Specific Purposes in Indonesia, 1(1), 12–17. <https://doi.org/10.33369/espindonesia.v1i1.23639>

- [56] Zamhari, A., Aina, Fitriani, Waliya, Septian, A., & Hadisatya, D. (2025). Language as Education for National Culture and Character. *Edukatif: Jurnal Kreativitas Dalam Pendidikan*, 6(1), 1–14.



M. Ridha Ajam is an academic with a doctorate in linguistics from the Faculty of Cultural Sciences, Khairun University. He has produced various linguistic studies that reflect the breadth of the field of linguistics. The specific areas of linguistic study he has written about include sociolinguistics, ethnolinguistics, and historical language relations. His work includes an analysis of colonial cultural acculturation and the absorption of foreign lexical elements in the communication of the Ternate community, a study of the communication style of Indonesian speakers in an intercultural context, and research on the historical relationship between the Taba and Patani languages. In addition, he has researched the conservation of the Ternate language through popular art media, the dynamics of oral traditions in Makian society, and the impact of social change on language practices in communities around mining areas, as well as students' perceptions of female teachers in the

context of language education and analysis of local policy management related to language use in social life. He can be contacted via email: mridhaajam@unkhair.ac.id



Mujahid Taha is a linguistics researcher at National Research and Innovation Agency of Indonesia. He is currently pursuing a doctoral degree in Linguistics. He has written numerous articles on linguistics, with a focus on applied linguistics. His research interest is ethnolinguistics. My email address and ORCID iD are: muja004@brin.go.id or jaisjais52@gmail.com and <https://orcid.org/0000-0002-9615-7349>.



Andi Sumar-Karman earned his Bachelor of Social Sciences in Anthropology from Hasanuddin University, Makassar, South Sulawesi, Indonesia, in 2001, and completed his Master of Arts (M.A.) in Anthropology at Gadjah Mada University, Yogyakarta, in 2013. Since 2021, he has been pursuing his Ph.D. in Anthropology from Hasanuddin University. Since 2006, he has served as a full-time lecturer at Khairun University, Ternate, North Maluku, Indonesia. He has actively published academic articles in many international journals focusing on his research interests that include ecological anthropology, particularly the dynamics of local communities and the environment. Sumar-Karman is also an active member of the Indonesian Anthropological Association (AAI). The author can be contacted via email: andisumarkarman@unkhair.ac.id; ORCID iD: 0009-0005-3538-6948.

The Bird Image in English and Arabic Poetry: A Stylistic Comparative Study of Selected Poems

Amjad Mohammad Almugheed
Jadara University, Jordan

Asma Hussein
University of Jordan, Jordan

Emad Abdallah Abuhammam
University of Jordan, Jordan

Ali Bani Ali
University of Jordan, Jordan

Abstract—The present comparative stylistic study explores the semiotic function and affinities of the avian imagery in selected poems from the poetic tradition of English Romanticism and Arabic verse, spanning classical to modern periods. The study employs the theoretical frameworks of Parallelism as defined by the American School of Comparative Literature and the Russian Formalist analytical methodologies to examine the cross-cultural literary convergences manifested in the selected poems. In addition, showing the significance of the bird makes people pleased and peaceful, especially with the other images of nature, such as spring, night, valleys, etc. Through close textual analysis of linguistic devices such as metaphor, personification, and symbolic transference, the study shows how avian imagery operates as a polysemous signifier, simultaneously encoding themes of transcendence, temporal liberation, and psychological interiority. Ultimately, the study demonstrates that despite the profound linguistic, cultural, and historical divergences, both traditions deploy ornithological imagery as a mediating discourse between human consciousness and natural phenomena. Consequently, these findings contribute to the broader theoretical discussions investigating universalist tendencies in literary expression while maintaining a sensitivity to the culturally specific hermeneutic contexts.

Index Terms—bird image, poetry, comparative study, parallelism, Formalism

I. INTRODUCTION

Natural appearances and symbols, including plants and animals, were reflected in the artistic image of poetry in various eras, starting from the Old Ages and ending with the modern eras. The presence of animals and birds, with all their characteristics and attributes, blew the imagination of poets, so they paid attention to the characteristics and behavior that characterize the world of animals, especially birds. In Arabic poetry, the poet attempts to convey his sentiments and what he wishes to tell the reader through his/her expression of the bird. He converses with the “bellows that follow him through the poem’s conclusion, and the poet uses the birds to embody a feeling that the recipient can relate to and try to express” (Nour, 2024, p. 78). This can be considered an attempt to convey feelings that poetry cannot express through other techniques like dialogue, mythmaking, and traditions. Moreover, Al-Qurashi (2023) states that in Old Arabic poetry, the bird image “helped people to understand the ancient lifestyles, religions, fables, social and historical conditions” (p. 56). Furthermore, the bird is aesthetically depicted as a symbol of freedom, hope, and happiness as well in various poems in Arabic literature.

Dambli (2019) concludes that “the dove has a large presence in pre-Islamic poetry because of its soul. It has a beautiful voice when it is singing, an essence of lamentation, so the pre-Islamic poet used it appropriately with these qualities” (p. 88). It can be noted that the bird has an intense presence from the Old and Modern Arabic ages. The Arabic poet needed to use the bird to form his artistic palette for many purposes. For example, when he was excited, he used it as “a threat to his opponent, and when describing battles, he used it to show strength” (Al-Hashimi, 2014, p. 77). Considering what was mentioned by some poets, it becomes clear that “they were immersed in the monologue of the birds. So, they showed them with their vast imagination as if he was able to speak and respond” (Nour, 2024, p. 78). Moreover, the poet became acquainted with the bird and took it as a friend, to whom he conveyed his concerns and secrets. However, this portrayal of the birds is similarly shown in different poetical forms of other nations, especially in the English tradition. The bird is shown similarly as a symbol of freedom and happiness on the one hand, and a shelter to some poets with their sadness on the other hand. However, D’coutho (2013) states that “birds are also a source of inspiration to poets to understand and relate with nature in different aspects like feelings, signs, beliefs, and as messages, more often even used as messengers to communicate to far and distant places” (p. 1014). Furthermore, poets have long utilized the natural world and its beauty

to convey their ideas and emotions. Springer (1990) comments on Whitman's use of the bird image as a "source of Romantic tradition and it evokes the whole bucolic style in English literature. The bird is a symbol of life and freedom" (p. 16). Birds were appreciated by most poets of English literature, especially the Romantic poets.

Ahmed (2021) argues that "birds are an important part of nature's beauty eco which is admired by the romantic poets; besides, birds are a source of inspiration to romantic poets in order to relate with nature in different aspects like feelings and beliefs" (p. 1802). On the other hand, birds can be used when they are sad, and they can be a relief to their souls. Birds also can be depicted in mythology. Warren (2018) states that "English poems of the bird are rich in allusions to the Greek story about its origin" (p. 411). However, the bird is depicted happily in many English poems, especially during the Romantic age. Birds can symbolize different feelings and sensations too. Abdallah et al. (2024) state that romantic poets "used birds as metaphors for life, the mind's marvels, forgiveness, hope, and despair. They can draw inspiration from birds to comprehend and connect with nature in various ways, such as through feelings, signs, beliefs, and messages" (p. 2793). Birds are often used as symbols of purity, innocence, and youth. Furthermore, many classical and modern English and Arabic poets identified the bird in their poems as a symbol of hope and freedom and others depicted it as a shelter for their sorrows and psychological dilemmas as analyzed in some poems selected in this study. The significance of this study is to focus on providing a stylistic comparative study between many poets from different linguistic, historical, religious, and geographical backgrounds. This study utilizes the depiction of the bird in poetry symbolizing its presence with humanitarian feelings and emotions.

II. RESEARCH METHODOLOGY

This study adopts analytical and comparative methods. English and Arabic selected poems will be compared and analyzed at different levels to assess and study their themes and ideas. In addition, this study relies on the conceptual and practical aspects of the American School of Comparative Literature. It uses the American School's Parallelism theory that focuses on "studying the historical and social evolution of literary texts" (Enani, 2005, p. 42). However, this study is also based on Russian Formalism's practical and theoretical aspects. It studies the linguistic techniques the poets used by analyzing "similes, metaphors, tones, symbols, and poetic styles...etc." (Bressler, 2015, p. 44). Moreover, this study analyzes each English and translated Arabic poem against one another showing similar images and themes stylistically. This study addresses these significant questions:

- 1- How do English and Arabic poets portray the bird in poetry?
- 2- What are the similar themes and linguistic devices used by these poets?
- 3- What does the bird symbolize in English and Arabic poetic traditions?
- 4- To what extent does the bird affect these poets' moods and thoughts?

III. DISCUSSION AND ANALYSIS

This study of linguistic parallelism starts with the portrayed and thematic images. First, the English poet S.T. Coleridge (1772-1834 CE) depicts the bird aesthetically in a personification. The speaker initiates the poem in an interrogative form "Do you know what the birds say? The sparrow, the dove," this double technique is used to catch the addressee's attention toward the bird. The speaker delivers two birds the "sparrow and the dove", he then continues to mention other birds like "the linnet and the thrush". These birds are saying "I love, I love", this indicates how the speaker symbolizes the bird as a sign of peace and love in life. All these birds get silent during winter in which the speaker highlights the winter's sorrow and loneliness. In the same line, the speaker mentions "so strong", he means the wind during the winter whereas nature is desperate and sorrowful. At the end of the first stanza, the speaker is wondering about the bird's speech, and then he recognizes it as a melodic song that the bird enjoys its words.

With hope, the speaker initiates the next stanza. He expresses the "green leaves", "blossom", and "sunny warm weather". This indicates how the speaker symbolizes the bird as a source of freedom, hope, and comfort. This bird sings with the elements of spring on a sunny and blossomed day:

Do you know what the birds say? The sparrow, the dove,
The linnet and thrush say "I love and I love!"

[...]

But green leaves, and blossoms, and sunny warm weather,
And singing, and loving, all come back together.
But the lark is so brimful of gladness and love,
(Coleridge, 1975, p. 101)

The speaker continues his comfortable and blessed mood due to the bird's song and recognizes the lark as glad and lovely. He mentions the word "love" many times in the poem to symbolize the bird as a sign of love and pleasure in nature. Moreover, spring with the bird's song bestows a sense of tranquility with nature. The greenness of the land with the image of the singing bird provides a happy tone to the speaker and the reader. This is what the other Romantic English poet William Wordsworth (1770-1850 CE) delivers in his "Lines Written in Early Spring", he presents the image of the birds in a personification whereas they "played and hoped" in the orchard during spring. The speaker shows how comfortable

he feels “While I sate reclined”, this indicates the image of mental peace with natural comfort. In the second stanza, the speaker enjoys seeing the birds playing and singing whereas he loves their existence in nature:

The birds around me hopped and played,
 Their thoughts I cannot measure: —
 But the least motion which they made
 It seemed a thrill of pleasure.
 (Doren, 1951, p. 51)

In another personification, these birds have certain ideas that the speaker cannot measure, but they indicate how pleased and tranquil these birds are for being in nature, especially during springtime. This view gives the speaker self-control and a positive psychological impact. In the end, the speaker feels grateful to nature, especially birds for keeping him happy and enjoying himself.

Similarly, in Arabic tradition, the classic Abbasid poet Abu-Tammam (803-845 CE) depicts a similar image of a singing bird on the tree's branch. While the speaker notices this bird inviting another “lunar” to enjoy its song. The speaker initiates the poem with the bird's glad mood in a personification in which it sings on the vibrating branches in spring. Metaphorically, the beloved lunar in the poem indicates the speaker's lover whereas he feels passionate and missed by her:

A songbird sang and made you passionate
 When he sang while the branches stretched
 Putting a stalk upon a stalk and called a lunar
 So, she shared his passion for hunting
 Two gathered in the shadow of the branches
 (Korkes & Mikhae'l, 1971, p. 34)

Nature's beauty and peace bless these two loving birds. The speaker indicates “morning” in the next lines as a metaphor to express the bird's happy mood while he is sitting with his beloved on the tree's branch. Morning during dewy weather makes these two birds enjoy nature's views and comfort. The speaker metaphorically blends the image of beautiful nature with love. During green nature, these birds share their ideas and emotions:

Oh, birds enjoy with a blessing
 In the morning, I am passionate
 Oh, for the departure, O son of Muhammad!
 Between the lovers is severe
 (Korkes & Mikhae'l, 1971, p. 34)

The speaker finishes up the poem by mixing the image of youth with the image of the sky. While the sky is bright and blue, the birds' souls are fresh and young. Moreover, the speaker indicates how this bird is generous and sensitive whereas other birds love to sing and twitter as well.

Another Arabic modern poet Ibrahim Touqan (1905-1941 CE) depicts the same image of the bird blending it with his feelings toward his beloved. In his poem “My Happiness When I See Her”, the speaker initiates it by mentioning his feeling toward his beloved. The speaker also resorts to the bird “Bulbul” to express his emotions objectively. He presents himself metaphorically as this singing bird that distributes enjoyment and love everywhere. In a simile, he is like the “bulbul” over the tree branches singing love and passion. The speaker personifies the “bulbul” coming and speaking to the speaker as a man who listens to his sorrows. This speaker is like the “bulbul” while they both sing and talk to each other. The speaker asks the “bulbul” after it listens to him crying to leave, but this bulbul is influenced by the kind of song that the speaker sings. Then, the speaker repeats to the bulbul how happy he feels when he sees his beloved “My happiness when I see her/ like the breeze of paradise”. Metaphorically, this “bulbul” is a symbol of joy and tranquility to the speaker:

My happiness when I see her
 Is my paradise which has a fire of passion
 A bulbul over the branches
 And my blessing is in my misery
 (Yousef, 2008, p. 55)

The speaker repeats his depiction with another image of the “bulbul,” a fantastic bird that attracts people to hear its beautiful sound. With compassion, this bulbul beautifully repeats his song to let the speaker enjoy it and stay in a suitable mood. In dialogue, this bird listened to the speaker's voice and asked him about his sorrow, but the speaker insisted that the bird leave him. Metaphorically, the speaker started to feel enjoyable and comfortable when he saw the bulbul. The image of the bulbul's wings and the presence of its feelings make the speaker peaceful too. In repetition, the speaker concludes the sentence “And my blessing is in my misery” to charge the bulbul to sing with his in the orchard. In the end, the speaker resorts another time to this bird to help him to forget his pains and dilemma of missing his beloved.

Inasmuch, similar images of the bird are portrayed in another poem by Wordsworth. In his “To the Cuckoo”, the speaker describes his peaceful mood in a green field listening to the cuckoo's sound around him. At the beginning of the poem, he declares that he hears this bird's songs and twittering in the garden. This joy is caused by the cuckoo's sound with the beautiful images and views of spring. Spring and the cuckoo make peoples' moods tranquil and joyful. In a personification,

the speaker calls the cuckoo, and, in a dialogue, he asks it to repeat his songs. The speaker blends the image of sunshine and spring with the charm of the bird to show how nature makes him happy while he is away from the material world:

O blithe New-comer! I have heard,
I hear thee and rejoice.
O Cuckoo! shall I call thee Bird,
Or but a wandering Voice?
(Curtis, 2009, p. 111)

The term “blithe new-comer” suggests that the cuckoo is unconcerned with the constraints of human existence. The bird’s joy is emphasized by the fact that it is unconcerned with the material concerns. The speaker is delighted as the cuckoo’s melodic call brings back happy memories. The stanza’s third and fourth lines metaphorically suggest that the speaker has only heard the bird’s call and has never seen it. Metaphorically, the speaker expresses doubt about the cuckoo’s actual identity by asking if it should be referred to as a bird or just a wandering voice. On the other hand, in the middle of the poem, the speaker rhetorically signifies that he is reminded of many incidents by the cuckoo’s lovely song, even if he just sings for the valley, sun, and flowers.

In the last stanza, the speaker repeats his admiration of the image of the cuckoo and describes it as “blessed” and happy:

O blessed Bird! the earth we pace
Again appears to be
An unsubstantial, faery place;
That is fit home for Thee!
(Curtis, 2009, p. 111)

Rhetorically, the speaker continues describing the earth as a place that looks like it belongs in a fairy tale. With the natural elements, the cuckoo returns to his green and charming home. The word “again” in the stanza shows how this bird admires his surrounding nature and home.

Similarly, the modern Arabic poet Ilia Abu-Madi (1889-1957 CE) depicts similar images of the bulbul when the “crow” listens to its song in the field. The speaker initiates the poem “The Crow Said” by comparing the crow’s and the bulbul’s wings and shape. The speaker personifies the crow as asking the bulbul about his fascinating wings. In the second line, the crow admires the singing bulbul and loves his style of twittering. Metaphorically, according to the crow’s perspective, the bulbul seems stronger, brighter, and more attractive. In a simile, the speaker shows the crow feeling himself like the bulbul. Then, this crow is wondering about the people’s praise towards it. In a paradox, the crow misses people’s speech and praise while they are drifting to the bulbul due to its beauty and perfectness:

The crows said a lot
And their infatuation with singing bulbuls
Why ears don’t hear about me like them
What is the difference between their wings and mine
(Abu-Madi, 1965, p. 80)

The speaker proceeds to show the “crow” questioning people why they love to hear and see the bulbul more than it. Symbolically, the bulbul presents the beauty and fascination of nature, especially the existence of the birds inside the speaker surrounding nature. In a personification, the bulbul feels luckier for being loved by people around. The speaker finishes the poem by recognizing the bulbul’s sound and song are more beautiful than its sound and presence.

Inasmuch, similar images of the birds are shown in different poems by other English and Arabic poets. For instance, the English romantic poet Percy Shelley (1792-1822 CE) portrays the image of the “skylark” singing and charming. In his “To A Skylark”, the speaker spots a skylark soaring overhead at the beginning of the poem. The melody is audible to him. The bird’s “unpremeditated” is lovely and spontaneous. Symbolically, the speaker is mesmerized by the bird’s melody and captivated by its flight as it disappears into the clouds. Although he can never see it, he is still able to hear it and feel its presence. The unadulterated joy that the speaker badly seeks is symbolized by the bird. In the next stanza, metaphorically, and according to the speaker, spring or “vernal” rains that pour on the flowers below, and the bird is a “high-born” girl who sings to her lover below. The skylark is the pinnacle of all “joyous” and resembles “rainbow clouds”. In the following stanza, the speaker is asking the skylark to reveal what inspires it to sing. The speaker states that the reason for the song is the beauty of the natural elements around both the skylark and the speaker:

Higher still and higher
From the earth thou springest
Like a cloud of fire;
The blue deep thou wingest,
And singing still dost soar, and soaring ever singest.
(George, 1991, p. 60)

In a simile, while the speaker is listening to the skylark’s song, this skylark runs and “floats” like the happiness that the speaker feels quickly when he listens to its song. Symbolically, this “skylark” is shown as the source of the speaker’s joy and comfort.

In another simile, in the middle of the poem, the speaker represents this skylark as a “poet” who hides himself to recite poetry and sing with a melody. This skylark and the poet make people sympathize and enjoy their songs. Another simile

is depicted in the following stanza, this skylark is like a “high-born maiden” who is in a “palace-tower”, this maiden is full of love and peace, especially with the sound of music.

Like other poets, the speaker in this poem mixes the image of the rain and spring with the image of the birds. Symbolically, the beauty of this bird is like the charming views of the raindrops on the grass and flowers, and the feeling of the land toward the rain is like the people's emotions when they hear this skylark's song:

Sound of vernal showers
On the twinkling grass,
Rain-awaken'd flowers,
All that ever was
Joyous, and clear, and fresh, thy music doth surpass.
(George, 1991, p. 60)

Metaphorically in the next stanza, the speaker pleads to this bird to “teach” him how to get “sweet thoughts” or feel pleased. The speaker symbolically sees this bird as a kind of love or “wine” that can make him joyful. This idea is repeated in the last stanza of the poem, the speaker highly appreciates the role of the skylark in making him happy, but he insists on the skylark to guide and “teach” him “gladness” and comfort:

Teach me half the gladness
That thy brain must know,
Such harmonious madness
From my lips would flow
The world should listen then, as I am listening now.
(George, 1991, p. 61)

Finally, the speaker explains that all people and the “world” should listen to the beautiful songs of this skylark. Metaphorically, these skylark songs indicate the happiness and enjoyment of people and nature.

Similarly, the Arabic poet Abu-Feras Al-Hamadani (932-968 CE) portrays the singing “dove” image. In his poem “I Say When a Dove Cries Near Me”, the speaker resorts to the dove to reduce his sorrows and pains. The speaker feels comfortable and enjoys it when he listens to its song. In the first lines, the speaker presents the image of the dove wailing and singing. He asks the dove about how he feels because he sees the dove full of sorrows and dilemmas as he has. He indicates his beloved and feels sad for being away from her. Metaphorically, the speaker is wondering about how sad this bird is while it is on a high tree branch. Rhetorically, the speaker describes this dove as his neighbor, and he cares about what it sings:

I Say When a Dove Cries Near Me
My neighbor, do you feel my situation?
Since love, I have never tasted comfort
And worries never crossed your mind
(A-Dahan, 1944, p. 50)

The whole poem is a metaphor for the speaker's resort to the dove to share his dilemma and pain. Symbolically, the dove is a source of tranquility and peace to the speaker. In the middle stanza, the speaker personifies the dove by showing it crying and singing, but he then feels comfortable being near the dove. However, the repetition of the line “Come, I need to share my worries with you, come” is represented to emphasize the speaker's sorrows and feelings. Finally, the speaker closes the poem by sharing the feeling of the dove with his and concludes how beautiful this dove is!

Similarly, this depiction is shown in another English poet's representation. Edgar Allan Poe (1809-1849 CE) represents a similar metaphor of the bird in his poem “The Raven”. In this poem, the speaker portrays a deep metaphor of how the raven is symbolized as a source of hope and tranquility. As in previous Arabic poet Abu-Feras Al-Hamadani's portrayal of the image of the dove and the beloved, in this poem, Poe tries to forget about his sorrows of remembering his beloved. Suddenly, when he was sitting alone the raven came near his window. This poem is made up of eighteen stanzas of six lines each. The speaker begins the poem by describing the gloomy setting that he was sitting in. It was “midnight dreary” whereas he was “weak and weary”. In the middle stanza, the “raven” comes near to him and talks in a personification. The speaker felt happiness and comfort when he remembered his lover:

Then this ebony bird beguiling my
Sad fancy into smiling,
By the grave and stern decorum of the
Countenance it wore,
“Though thy crest be shorn and shaven, thou,”
I said, “art sure no craven,
(Poe, 2013, p. 10)

The speaker sees that this raven is like an angel who came to reduce his sorrows and pains. The speaker asks the raven whether he is a bird or “beast” upon the “sculptured bust above his chamber door”. The speaker admires the raven's presence around him. In another personification, the speaker depicts the bird sitting lonely and speaking and repeating the only word “Nevermore” when he asks it about itself. However, in the end, the raven keeps sitting in front of the speaker and makes his mood comfortable and more peaceful because of the raven's natural beauty.

Similarly, in Arabic poetry, the modern Arabic poet Nazik Al-Mala'ka (1923-2007 CE) depicts similar images of the bird as the previous Arabic and English poets represent. In her "In the Arms of Nature", the speaker returns to the bird to enjoy nature. He depicts it as a symbol of beauty and perfectness. Metaphorically, the speaker invites people to dream of the birds and imagine their charm. Then, he obliges his readers to look at the birds and enjoy their singing and flight. In addition, the speaker represents the image of the "lunar" in the poem with the image of spring and the purl. The image of the singing bird in the green field makes the speaker's mood happy and peaceful "And dream of birds in the shade of wood/ Between flying and tweeting" (Al-Basri, 1971, p. 30). However, the speaker invites the listeners of the "lunar" to hear her song. Her song symbolizes people's happiness and excitement, especially when they listen to it. These images of the birds are repeated in other poems of Al-Mal'ka, in her "Birds Water the Existence Cups", the speaker also delivers the image of the bird as a symbol of hope, pleasure, and life. Moreover, the speaker in this poem describes the bird as a symbol of freedom, especially when he listens to a bird singing about its liberty in a cage. In a personification, the speaker added human qualities to the birds, as they "water" so he turned the birds into a person who waters Existence is a cup of constant, sweet songs that are also poured into the ear of the green hummocks and hills, the birds are always tweeting sweet sounds that inspire existence that listens to it. Moreover, the birds are happy, singing those songs above the great trees, happy and tranquil, but the speaker below calls it as if it were addressing a rational person, urging it to fly away from this world which he describes as painful, arousing grief and sadness. In his view, the bird would be happy if it was flying in the sky:

And birds watering existence with
Cups of their pure songs of torment
O truthful one who tweets above the noise
Flee from this painful existence
(Al-Basri, 1971, p. 41)

Metaphorically, the speaker recognizes that the bird's song is the psychological therapy that makes him pleased while he enjoys nature and its beautiful influence on his mind. On the other hand, a similar image is depicted by Al-Mal'ka in another poem. In her "In the Face of the Bird", the speaker portrays the image of the bird as sad at first, because its nest was destroyed. Metaphorically, the bird has nothing but to cry. The bird lost its nest, so it cried in sadness for its shelter, but it stayed nearby. Also, it waited a long time, hoping that he might get it back again. The poet cried in grief over what had happened. Rhetorically, its homeland is in ruins, but it has no choice but to cry and grieve over it, deviating from her imagination of being sent. Then, this bird started to sing whereas the speaker symbolically considers it as a hopeful motif in his life. The speaker closes the poem by concluding that this bird is a symbol of patience, hope, and happiness. However, this image of the bird is shown in an English romantic poem by John Keats (1795-1821 CE) "Ode to a Nightingale". This poem centers on a speaker who is standing in a pitch-black woodland and enjoying the nightingale's lovely and captivating song. Metaphorically, this inspires a deep and meandering thought by the speaker on nature and the beauty of the birds. Symbolically, the speaker finds solace in the nightingale's song. The speaker concludes to be alone, but when he hears the "nightingale's" song, he starts to feel pleasurable and hopeful. The speaker confesses that this bird is his resort to be away from the sorrows of the material world. In a paradox, the speaker feels at the beginning of the poem lonely and sad, but at the end of the poem, the speaker starts to feel with a unique creature the "nightingale" and wishes it could not fly or die:

Thou wast not born for death, immortal Bird!
No hungry generations tread thee down;
The voice I hear this passing night was heard
In ancient days by emperor and clown:
(Stillinger, 1982, p. 22)

Metaphorically, because of the beautiful and soft existence of the "nightingale", the speaker at the end of the poem is wondering if this bird's song was "a vision or awakening dream". In addition, the bird's song is a symbol of hope and pleasure that the speaker enjoys and dreams to hear again.

IV. CONCLUSION

This study invokes the depiction of the bird image in poetry through a thorough analysis. It studies and compares the selected poems showing the images of the bird emotionally, stylistically, and thematically. The bird in English and Arabic poetry is similarly depicted as a source of gladness, freedom, and peace. Moreover, these poets shed light on different natural elements, such as spring, animals, mornings...etc. blending their charm with the bird's twittering and songs. Many natural and humanitarian occasions are sung on the birds' tongues including exciting and hopeful tracks. In these selected poems, the speakers enjoy the presence of the birds in nature, and they depict them in a lonely situation to show the readers how the birds help them to cope with their alienation and sadness. Finally, the birds are portrayed as a symbol of enjoyment and joy too.

REFERENCES

- [1] Abdallah, N., Ali, E., Begum, T., & Eltom, S. (2024). The Symbolic Significance of Using Birds in English Romantic Poetry: "Ode to a Nightingale by John Keats" as a case study. *Pakistan Journal of Life and Social Sciences*, 22(2), 2787-2796.
- [2] Abu-Madi, Ilia. (1965). *Poems of Ilia Abu-Madi*. Lebanon; Dar Al-Awda for Publication.
- [3] Ahmed, Awadia. (2021). Effects of Symbolism of Birds in British Romantic Poetry. *European Academic Research*, 5(3), 1801-1813.
- [4] Al-Basri, AbdelJabar. (1971). *Nazik Al-Mala'ka: poetry and theory*. Baghdad: Dar Al-Huria for Publication.
- [5] Al-Dahan, Sami. (1944). *Poems of Abu-Feras Al-Hamadani*. Beirut: The French Institute.
- [6] Al-Hashimi, Hamad. (2014). *Addressing The Bird in The Pre-Islamic Poetry Till the End of the Fifth Century*. KSA: Um Al-Qura University Press.
- [7] Al-Tarifi, Yousef. (2008). *Ibrahim Touqan, his life and poetry*. Jordan: Al-Ahlya for Publication.
- [8] Al-Qurashi, Amani. (2023). Extrapolation of the image of the motherhood of the bird in Al-Qama Al-Fahl's Diwan. *Journal of Arabic Language and Literature*, 2(4), 54-66.
- [9] Awad, Korkes, & Awad, Mikhae'l. (1971). *Abu-Tammam Al-ta'I, his life and poetry*. Baghdad: Al-Irshad Press.
- [10] Bressler, C. (Ed.). (2015). *An introduction to theory and practice* (5th ed.). Library of Congress.
- [11] Coleridge, Ernest. (1975). *The complete poetical works of S.T. Coleridge*. UK: Oxford University Press.
- [12] Curtis, Jared. (2009). *The poems of William Wordsworth*. UK: Humanities Ebooks.
- [13] Dambli, Mamado. (2019) The use of birds in the pre-Islamic poetry Antara Ibn Chaddad as a model. *Revue Annales du Patrimoine*, 19(1), 83-98.
- [14] D'couth, Roweena. (2013). The Role of Birds in English Literature and Poetry. *International Journal of Science and Research*, 5(4), 1014-1016.
- [15] Doren, Mark. (1951). *The works of William Wordsworth with an introduction*. New York: Random House.
- [16] Enani, Mohamad. (2005). *Theories of Comparative Literature*. Egypt: Cairo Publications.
- [17] Nour, Afaf. (2024). The Symbolism of the Bird in Arabic Poetry. *Journal of Al-Basra Studies*, 19(51), 87-98.
- [18] Poe, Edgar. (2013). *Edgar Allan Poe: Tales and poems*. UK: Maplewood Books.
- [19] Stillinger, Jack. (1982). *John Keats, complete poems*. Harvard University Press.
- [20] Springer, Nancy. (1990). *The Source of Bird Imagery in The Poetry of Walt Whitman*. USA: Florida Atlantic University.
- [21] Warren, Michel. (2018). *Birds in Medieval English poetry: Metaphors, realities, transformations*. UK: Boydell & Brewer.
- [22] Woodberry, George. (1991). *Complete poetical works of Percy Shelley*. USA: Boston and New York Houghton Mifflin Company.

Amjad Almuqheed is a lecturer of Russian Language at different universities in Jordan. He received his Ph.D from Oles Honchar Dnipropetrovsk National University, Ukraine, in 2022. His publications focus on Translation and Comparative Philology and Phonetics. Moreover, he is interested in linguistic studies and teaching Russian language to non-speakers.



Asma Hussein is currently a full-time lecturer at the University of Jordan. She was an Assistant Professor at Jerash University/Jordan. She was an Assistant Professor at Al-Ahliyya Amman University/Jordan (October 2024-September 2025). She was a lecturer at Georg August University, Göttingen, Germany (2021-2023) where she taught courses on the MA program and a teaching assistant at Ada und Theodor Lessing Vhs Hannover, Germany (2018-2020). She received her Ph.D. in English Language and Literature in 2017 from Masaryk University (Czech), where she was a teaching assistant at the department of English and American studies (2012-2017). Her publications focus on literary criticism and Post-colonial Studies.



Emad Abdallah Abuhammam is currently a lecturer at the University of Jordan, Department of English. He received his Ph.D in Comparative Literature in poetry from the University of Seville, Spain, in 2024. He also completed his MA in English Literature from Mutah University-Karak-Jordan/2021, and a bachelor's degree in English Language and Literature from Tafila Technical University-Tafila-Jordan/2016. Dr. Abuhammam has published many articles in reputable journals on poetry. He is interested in comparative studies in poetry.

Ali Bani Ali is a lecturer at the University of Jordan. He works at the Foreign Languages Faculty, Department of Asian Languages. He graduated from Turkey in 2021, and he is interested in comparative, translational literary studies of the Turkish language.

Appraisal Framework Instruction and Its Impact on English Writing: A Study Among Vietnamese University Students

Hong B. Nguyen*

Faculty of English, Thuongmai University, Hanoi, Vietnam

Abstract—This study aims to investigate the effectiveness of applying the Appraisal framework, derived from Systemic Functional Linguistics, to improve English writing skills for Vietnamese university students. The Appraisal framework, developed by Martin and White (2005), provides a tool for analyzing lexical resources to help learners understand and apply expressions of Attitude, Engagement, and Graduation in texts. This study uses a quasi-experimental design, pre-testing and post-testing with 60 university students in Vietnam. The experimental group received explicit instruction on the Appraisal resources (Attitude, Graduation) integrated into a writing course for 12 weeks, while the control group followed a regular writing syllabus applied for all students at the university. Results indicated that the experimental group showed statistically significant improvements in more appropriate use of Attitude markers and a wider range of Graduation resources compared to the control group. This consequently helped to improve their writing performance as a whole. Student feedback suggested that the Appraisal instruction enhance their metalinguistic awareness and provide practical tools for producing more evaluative texts, despite some initial challenges with terminology. The pedagogical potential of the Appraisal framework in the context of EFL writing instruction in Vietnam was highlighted to develop students' evaluative and interactive competencies in English writing.

Index Terms—evaluative language, writing, Appraisal framework, Attitude, Graduation

I. INTRODUCTION

Effective writing skills are essential for academic and professional achievement in our increasingly globalized world. Nonetheless, Vietnamese learners learning English as a Foreign Language (EFL) encounter considerable obstacles in advancing these skills (Duy & Trang, 2022; Nghi & Truong, 2023; Phuong, 2021; Tran, 2023). In addition to common challenges in vocabulary, grammatical structures, organization of ideas and background knowledge, an important but often overlooked aspect is the ability to express the writer's attitude appropriately. In traditional Vietnamese language training, grammar and vocabulary acquisition is given more weight than the development of higher-order rhetorical and interpersonal skills, which are crucial for creating sophisticated and persuasive writing. In particular, Vietnamese EFL students' writing frequently lacks the capacity to interest the reader through the strategic use of language and to communicate attitude, including assessments, judgments, and emotional reactions. As a result, essays may become inexpressive and fail to convey the desired nuance of meaning, leading to ineffective communication and a lack of connection with the reader. These limitations can hinder students' performance and career prospects in contexts where effective communication in English is essential.

Recognizing the importance of evaluative language in academic writing, a number of studies have used the Appraisal framework (Martin & White, 2005) to examine how evaluative and interpersonal meanings are constructed in particular genres and discourses such as academic articles (Hood, 2004, 2010; Hyland, 2005), student writing (Coffin, 2003; Swain, 2010), and narrative and historical texts (Coffin, 2006; Macken-Horarik, 2003). More relevant to the present research are investigations into the pedagogical possibilities of the Appraisal framework in writing instruction. According to these studies (Lee, 2015; Schleppegrell, 2004), learners' metalinguistic understanding of evaluative language may be improved by examining model texts via an Appraisal lens. Furthermore, studies in L1 and L2/EFL contexts have demonstrated that specific aspects of student writing, including stance-taking, argumentation, critical evaluation, and authorial voice, can be improved through explicit instruction based on the Appraisal framework (Coffin & O'Halloran, 2005; Derewianka, 2007; Wu & Allison, 2007; Yang, 2016). These studies frequently emphasize that providing students with the analytical skills and vocabulary to discuss interpersonal language choices helps them better control these elements in their own writing. Although the Appraisal framework is often applied to complex academic discourse analysis, its potential in supporting students at an intermediate level of English to improve their ability to express attitudes in simple English texts is still underexplored. Equipping students with a basic understanding of attitudinal language choices may help them write more confidently, expressively and communicate more effectively even at this level.

* Corresponding Author.

In Vietnam, some studies have explored the function of evaluative language in specific genres (Nguyen, 2021; Nguyễn & Nguyễn, 2020). Nevertheless, further study is still required to examine the effectiveness of the Appraisal framework as a pedagogical tool in Vietnamese EFL classes. This study attempts to fill this gap by investigating how writing abilities of EFL students at a university in Vietnam may be improved through the instruction of the Appraisal framework in general and the Attitude and Graduation systems in particular. Findings of this study have the potential to make contributions to the field of applied linguistics, EFL pedagogy, and, ultimately, to the development of Vietnamese EFL learners as more confident and proficient writers of English.

II. LITERATURE REVIEW

A. The Appraisal Framework

This study is based on the Appraisal framework (Martin & White, 2005), which is part of the school of Systemic Functional Linguistics (SFL) (Halliday & Matthiessen, 2014). SFL views language as a social semiotic system, which means language is both a resource for meaning-making and a product of social context. Accordingly, language is proposed to fulfil three metafunctions: (a) Ideational metafunction; (b) Interpersonal metafunction; and (c) Textual metafunction.

In order to analyze how language users communicate assessment, commitment, and alignment, Martin and White (2005) systematized the Appraisal framework, which goes further into the interpersonal metafunction (Figure 1). This framework comprises three systems: Attitude, Engagement, and Graduation. Attitude addresses how emotions, opinions, and assessments are expressed. Engagement examines how speakers and writers present themselves and their audience in relation to the ideas they are promoting. Graduation deals with the scaling of attitudinal meanings. Within the scope of this study, systems of Attitude and Graduation are the main focus and, therefore, presented in more detail.

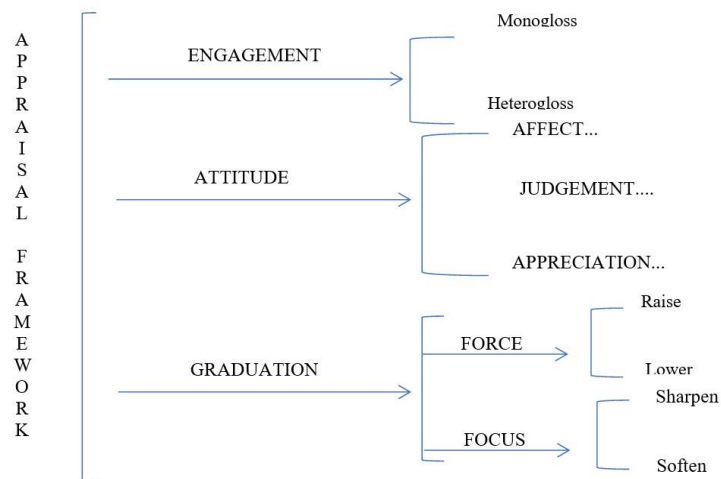


Figure 1. An Overview of the Appraisal Framework (Martin & White, 2005, p. 38)

(a). Attitude System

As a system within the Appraisal framework, Attitude focuses on how speakers or writers express feelings, judgements, or evaluations about people, things, events, or phenomena. Attitude is divided into three main subsystems: Affect, Judgement, and Appreciation (Figure 2).

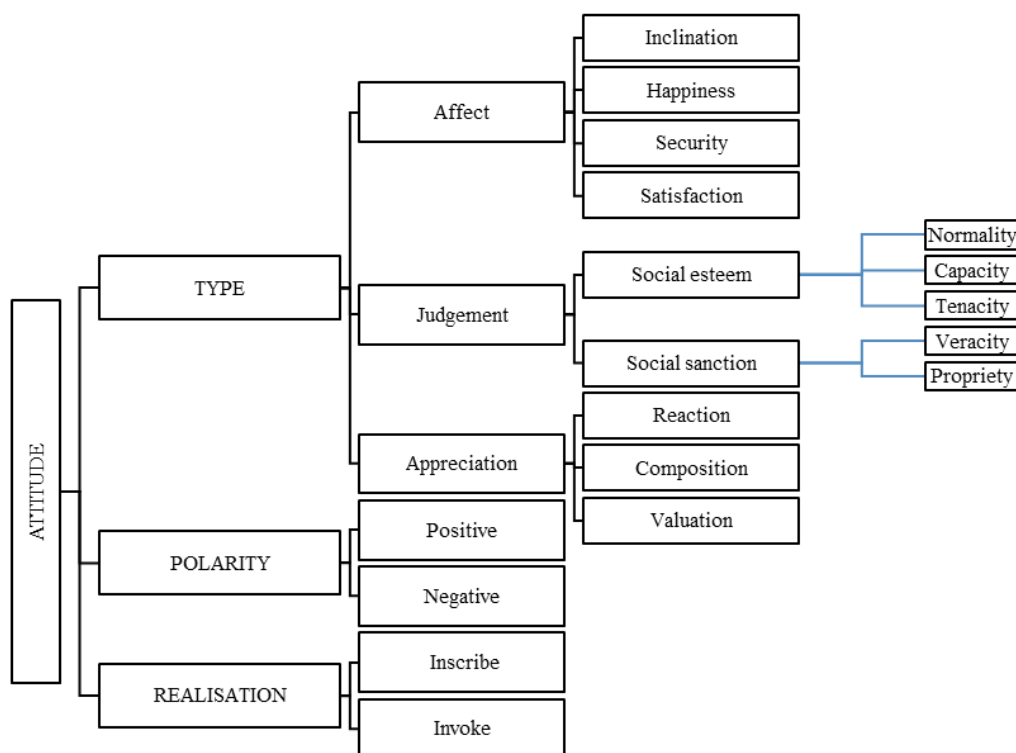


Figure 2. The Attitude System (Martin & White, 2005)

Affect consists of lexical resources for emotional states or emotional reactions to an event, situation, or object. Affect can be classified based on polarity and four emotion pairs (Dis/Inclination, Un/Happiness, In/Security, and Dis/Satisfaction.)

Judgement refers to attitudes toward people and how they behave. Judgement is divided into two main categories: Social esteem and Social sanction.

Social esteem is judged from the following aspects: (a) Normality: how special/different someone is; (b) Capacity: how capable someone is; and (c) Tenacity: how persistent someone is.

Social sanction evaluates people through two factors: (a) Veracity: evaluate someone's level of honesty/trustworthiness; (b) Propriety: evaluate someone's morality based on common values recognized by society.

Appreciation refers to evaluations of things. Appreciation includes three categories: Reaction, Composition and Valuation. Reaction shows the impact of the object being evaluated on the evaluator. Composition is the assessment of the structure of things and phenomena in two aspects: Balance and Complexity. Valuation is the assessment of the value of things and phenomena, answering the question of whether they are valuable or not in terms of innovation, authenticity, timeliness, etc.

(b). Graduation System

Both the Attitude and Engagement systems share the characteristic of gradability. In the Attitude system, gradability refers to the degree of positivity/negativity of evaluations, while in the Engagement system, it refers to the degree of involvement of the speaker/writer in the utterance being made. Therefore, Martin and White (2005) emphasize that Graduation is “the center of the overall system” (p. 136) and Attitude and Engagement are two aspects of Graduation with different natures of the meaning being graded.

Graduation is divided into two subsystems based on the category being evaluated: Force and Focus (Figure 3).

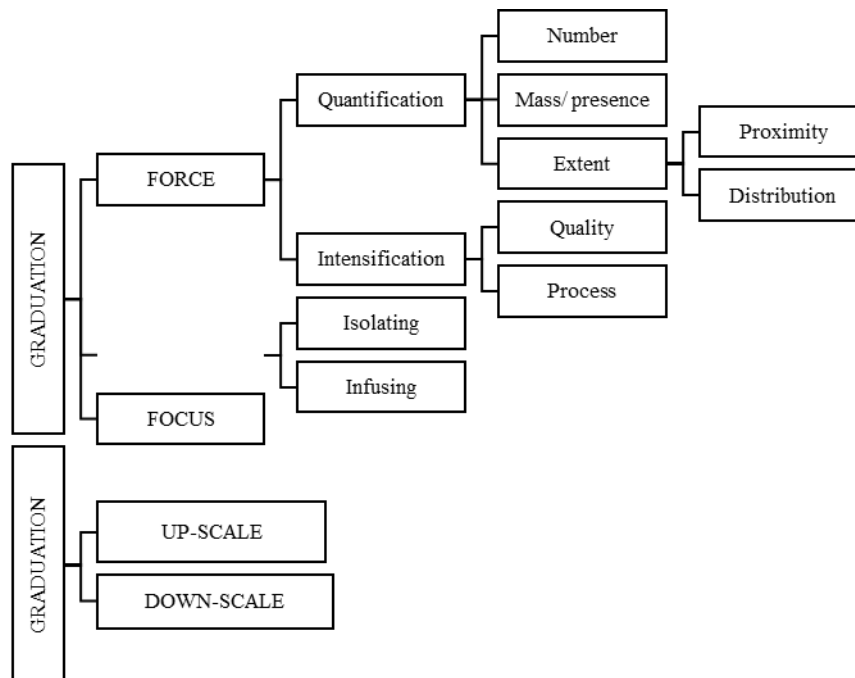


Figure 3. The Graduation System (Martin & White, 2005, p. 154)

Force includes Intensification and Quantification. Intensification refers to the grading of quality and process. Quantification represents an imprecise estimate of quantity, shape, size, and extent of time and space, measured in terms of proximity or distribution.

Focus is mainly applied to categories that are not inherently gradable. Accordingly, Focus is an expression to confirm the degree to which the evaluated object matches the assumed origin such as: upgrading to specify and affirm the characteristics of the prototype or downgrading to describe an object having a peripheral relationship with the category.

The Appraisal framework is of great importance in discourse analysis and linguistics because it provides a systematic tool for deciphering the social values, emotions, and attitudes expressed through language. It helps researchers understand how language is used to construct and maintain social relationships, regulate consensus or debate, and express personal commitments in communication. The framework also supports the analysis of different genres of texts such as journalism, literature, academic essays, and legal texts, contributing to clarifying the role of language in the social construction of meaning.

B. The Appraisal Framework in Writing Analysis and Pedagogy

The Appraisal framework has been extensively applied by scholars worldwide (Coffin, 2003; Hood, 2004, 2010; Hyland, 2005; Macken-Horarik, 2003) to examine a variety of genres, such as student writing, academic papers, media texts, and narratives. Hood (2004) highlighted the importance of Appraisal in explaining the multifunctional roles of texts (especially research proposals), including speaker positioning, politeness expression, and argument construction. Coffin (2003) employed the Appraisal framework to understand how media stories shape reader positioning through the accumulation of evaluative values, hence impacting readers' perceptions and emotions. Meanwhile, Macken-Horarik (2003) concentrated on the function of Appraisal in narratives, examining how the framework's linguistic resources enable readers to empathize with characters and assess their morality while also strengthening the bond between the author and the audience. These studies show that the Appraisal framework is a useful linguistic tool for revealing interpersonal relationships and evaluative aspects in a variety of text genres.

In educational contexts, a number of studies have demonstrated the effectiveness of applying Appraisal analysis in teaching writing skills, particularly in raising students' awareness of evaluative language and enhancing their ability to express stance. Schleppegrell (2004) showed how students can improve their writing skills by identifying and understanding the linguistic resources to express feelings, evaluations, and personal opinions in writing. Lee (2015) used the Appraisal framework to explore how university students express evaluative stances in persuasive essays. The results revealed that students predominantly relied on Reaction, with limited range and intensity, and rarely employed Valuation or Graduation. As a result, their arguments were often simple and unpersuasive, highlighting the need for explicit instruction on the strategic use of evaluative language in academic writing. Other studies also demonstrated positive effects of the Appraisal framework on certain aspects in not only first language (L1) but also second language (L2) contexts, including stance-taking, argumentation, critical evaluation, and the development of authorial voice (Alhuthali, 2024; Coffin & O'Halloran, 2005; Liu, 2013; Swain, 2010; Wu & Allison, 2007). For instance, Alhuthali (2024) used a qualitative discourse analysis to investigate how Attitude, Engagement, and Graduation resources were distributed across two EFL writing genres. Findings showed that genre significantly influenced students' evaluative

choices. Argumentative texts contained more Judgement and Graduation resources to enhance persuasion, whereas descriptive texts relied more on Appreciation, highlighting the role of genre awareness in developing evaluative competence in EFL writing. This study suggested tailored instructional approaches be necessary to develop students' evaluative writing skills. Liu (2013) examined how Chinese EFL university students express evaluation in English argumentative essays. Through qualitative and quantitative discourse analysis, the study found that students tended to rely heavily on Affect and Judgement while underusing Graduation and Engagement, resulting in limited rhetorical sophistication. Based on the results, Liu emphasized that explicit instruction on evaluative language can help EFL learners develop stronger stance-taking and more persuasive academic writing.

In summary, the Appraisal framework can be said to be not only an effective linguistic tool for deeply understanding evaluative dimensions in texts but also a valuable pedagogical resource that helps students improve their writing abilities, especially in expressing stance and engaging readers in a flexible and subtle way.

III. METHODS

A. Research Design

This study applied a quasi-experimental design because of its practicality in educational settings, where random assignment of students to classes is often not feasible. However, this design still allows comparison of changes over time between the group receiving the intervention and the group not receiving the intervention.

Participants were 60 second-year students at Thuongmai university in Vietnam. Two intact classes were selected based on criteria of similarity in entry English proficiency (determined by previous course grades) and class schedule. One class was assigned as the experimental group (EG, $n = 30$) and the other as the control group (CG, $n = 30$). Two classes were taught by the same teacher. The general English proficiency of students was determined at level B1+ according to the Common European Framework of Reference CEFR.

B. Procedures

The study lasted for 12 weeks. In the first week, both the EG and CG groups took a pre-test. From weeks 2 to 11, the intervention phase took place. The experimental group received one hour of instruction per week, in which the concepts and application of the Appraisal framework were integrated into the regular writing syllabus (Appendix A). Activities include explicit explanations of Attitude and Graduation systems using simplified language and concrete examples; analysis of sample essays from an Appraisal perspective; practice exercises from sentence to paragraph level to reinforce the use of Appraisal resources; peer feedback activities between students using Appraisal-based criteria; and guided writing and revising essays consciously using the Appraisal framework. Supporting materials included theory summaries, checklists, sample essays, and questions.

During the same period, the control group attended writing lessons aligned with the general program required to all students at the university, using the coursebook *Life – Intermediate level* (published by National Geographic Learning). In the final week (week 12), both groups took a post-test on the same topic as the pre-test. To ensure the equivalence for comparison between the two groups (EG and CG), students were given clear instructions to compose their writing (about 200 words) in 30 minutes under similar controlled conditions.

Finally, the EG group completed a questionnaire. The questionnaire included Likert-type questions and open-ended questions, focusing on students' perceptions of the usefulness of the Appraisal concepts, their application to writing, difficulties encountered, and reflections on the intervention period.

C. Data Analysis

The collected data were processed using both quantitative and qualitative methods. First, the pre-and post-tests were analyzed carefully to identify and categorize realizations of evaluative language (lexical resources for Attitude and Graduation) employed. The analysis was conducted using a manual coding method by an expert in the Appraisal framework, who has five years of experience in this field. The overall quality of the writing was also assessed using an analytic rubric including 4 specific criteria (Content, Organization, Vocabulary and Grammar) with scores (1-10) (Appendix B).

Then, quantitative analysis, performed using SPSS version 26, was carried out, including calculating the frequency of occurrence (per 100 words) of the realizations of linguistic resources indicating Attitude and Graduation in the pre- and post-intervention essays of both groups. Statistical tests conducted were: (1) an inter-rater reliability test (Intraclass Correlation Coefficient – ICC) to ensure consistency between the two raters in scoring students' writing across all criteria; (2) independent samples t-test to compare the initial similarity between the two groups on the pre-test; (3) multivariate analysis of variance (MANOVA) to examine differences between the two groups on multiple categories of Attitude and Graduation; (4) analysis of covariance (ANCOVA) to compare the post-test results between the two groups after controlling for the effect of pre-test scores; and (5) Pearson correlation analyses to investigate the relationships between the use of evaluative language (in terms of Attitude and Graduation) and the post-test writing performance of the experimental group. The level of statistical significance (alpha) was set at $p < .05$. Likert data from the questionnaire were summarized using descriptive statistics (frequencies, percentages). Qualitative results were used to supplement, enrich, and further explain the quantitative results.

IV. RESULTS

A. Inter-Rater Reliability

To ensure the consistency and objectivity of the scoring process, the writing tests were independently rated by two qualified raters using a rubric consisting of 4 criteria (Content, Organization, Vocabulary, and Grammar). Inter-rater reliability was examined for both the pre-test and post-test scores using the Intraclass Correlation Coefficient (ICC). The results indicated a high level of agreement between the two raters across all scoring categories (Table 1).

TABLE 1
INTER-RATER RELIABILITY (ICC) OF RUBRIC CRITERIA AND TOTAL SCORES

Criterion	Pre-test				Post-test			
	Intraclass correlation	95% Confidence interval		<i>p</i>	Intraclass correlation	95% Confidence interval		<i>p</i>
		Lower bound	Upper bound			Lower bound	Upper bound	
Content	.860	.766	.917	<.001	.882	.802	.929	<.001
Organization	.834	.721	.901	<.001	.768	.612	.862	<.001
Vocabulary	.850	.749	.911	<.001	.819	.696	.892	<.001
Grammar	.764	.604	.859	<.001	.811	.683	.887	<.001
Overall score	.929	.881	.958	<.001	.912	.852	.947	<.001

For the pre-test, ICC values were .86 for Content, .83 for Organization, .85 for Vocabulary, .76 for Grammar, and .92 for Overall writing. For the post-test, ICC values remained consistently high, with .88 for Content, .76 for Organization, .81 for Vocabulary, .81 for Grammar, and .91 for Overall writing. According to Koo and Li's (2016) benchmarks, these coefficients indicate good to excellent inter-rater reliability ($ICC > .76$) at both testing stages, confirming that the scoring process was consistent and reliable throughout the study, ensuring the validity of the writing assessment results.

B. Initial Equivalence of Groups

Before the intervention process began, independent samples t-tests comparing EG and CG groups on pre-test writing scores and frequencies of Attitude and Graduation realizations had been carried out. According to reports, there were no statistically significant differences between the experimental and control groups on overall writing scores or the frequencies of Attitude and Graduation resources in the pre-test (Table 2).

TABLE 2
GROUP STATISTICS AND INDEPENDENT SAMPLES T-TESTS OF THE PRE-TEST

Pre-test statistics	EG (<i>M</i> ± <i>SD</i>)	CG (<i>M</i> ± <i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i> -value
Pre-test scores	6.4583 ± .46463	6.5043 ± .63044	<i>t</i> (53.326) = -.322	.749
Attitude resources	8.6343 ± 2.58501	8.8797 ± 2.53291	<i>t</i> (58) = -.371	.712
Graduation resources	3.0673 ± 1.29684	3.2257 ± 1.92665	<i>t</i> (58) = -.373	.710

Specifically, the mean scores of EG ($M = 6.45$, $SD = .46$) and CG ($M = 6.50$, $SD = .63$) were not significantly different, $t(53.326) = -.322$, $p = .749$. Therefore, there is no statistical evidence to suggest that the two groups differ in terms of their mean scores.

As for frequencies of evaluative resources employed in the pre-test, reports showed similar results. The average frequency of Attitudinal resources employed by the EG and CG groups was 8.63 and 8.87 per 100 words, respectively. The average frequency of Graduation resources was much lower in comparison with Attitude, with $M = 3.06$ among EG students and $M = 3.22$ among CG students. However, the overall Attitude frequency ($t(58) = -.371$, $p = .712$) and Graduation frequency ($t(58) = -.373$, $p = .710$) both showed $p > .05$. These suggest the two groups were reasonably comparable at the outset of the study.

C. Post-Test Results

(a). Effects on Attitude Resources

A multivariate analysis of variance (MANOVA) was conducted to examine the effect of the intervention process on students' use of Attitudinal resources. Results (Table 3) revealed a statistically significant overall difference between the EG and CG groups on the combined dependent variables, $Wilks' \Lambda = .57$, $F(4,55) = 9.992$, $p < .001$. Follow-up analyses indicated significant differences in frequencies of all categories: Affect ($F(1,58) = 8.479$, $p = .005$, partial $\eta^2 = .128$), Judgement ($F(1,58) = 12.579$, $p = .001$, partial $\eta^2 = .178$), Appreciation ($F(1,58) = 28.197$, $p < .001$, partial $\eta^2 = .327$) and the totality of all Attitudinal resources ($F(1,58) = 22.693$, $p < .001$, partial $\eta^2 = .281$).

TABLE 3
MANOVA RESULTS FOR ATTITUDE FREQUENCIES (PER 100 WORDS) AT POST-TEST

Category	Group	Mean $\pm$ SD	$F(1,58)$	p -value	Partial η^2
Overall Attitude	EG	11.0923 $\pm$ 3.20697	22.693	<.001	.281
	CG	7.9410 $\pm$ 1.68641			
Affect	EG	1.9333 $\pm$ 1.29247	8.479	.005	.128
	CG	.9807 $\pm$ 1.24116			
Judgement	EG	.8570 $\pm$.83523	12.579	.001	.178
	CG	1.8583 $\pm$ 1.30138			
Appreciation	EG	8.3027 $\pm$ 2.66513	28.197	<.001	.327
	CG	5.1017 $\pm$ 1.94907			

These results suggest that the intervention had a significant positive impact on students' employment of Attitudinal realizations. In general, intervened students use much more emotional expressions in their writing than controlled ones after the intervention period. However, it is noteworthy that CG students particularly employed more Judgement resources than EG students.

Qualitative analysis of students' writing indicated a significant increase in frequency and variety of realizations of Attitudinal resources between the pre-test and post-test among students in the experimental group (Table 4).

As can be seen from Table 4, the frequency of all Attitudinal resources in the post-test was 11.09%, an increase of 2.46 percentage points compared to the pre-test result (8.63%). Concerning Attitude's sub-types and categories, it can be seen that students' use of negative evaluation and Judgement of human behaviors did not show obvious changes between the two tests, or even slightly decreased. This may be explained by the fact that after being introduced about the Appraisal framework and becoming aware of how to use evaluative language, students became more conscious and cautious when employing Judgement to assess people's behaviors in writing. Meanwhile, those who were not exposed to the framework still tended to make behavior judgements more freely and, therefore, more frequently. However, the employment of Affect and Appreciation reported substantial upward changes. The frequency of Affect dramatically increased from 1.09% to 1.93% whereas that of Appreciation was 6.65% and 8.30%, respectively. Furthermore, positive emotional expressions tended to appear more frequently in the post-test writing.

TABLE 4
ATTITUDE FREQUENCIES IN THE PRE- AND POST-TESTS AMONG EG STUDENTS (PER 100 WORDS)

EG group	Polarity		Sub-types of Attitude			Total
	+	-	Affect	Judgement	Appreciation	
Pre-test	8.31	0.33	1.09	0.89	6.55	8.63
Post-test	10.69	0.40	1.93	0.86	8.30	11.09

Realizations for these emotions also became a lot more diverse after the intervention. For example, to express Inclination, students used *want, hope, wish, desire*, etc. (e.g., *I want to escape from the hustle and bustle of work and study*); Happiness was expressed via such words as *enjoy, excited, prefer, enthusiastic, passionate*, etc. (e.g., *Visitors are also extremely **excited** about tasty local cuisine such as Pho, bun dau mam tom, com lang Vong, etc.*); realizations of Security included *familiarity, confident, sure, safe*, etc. (e.g., *However, what makes me love my homeland is still the feeling of **familiarity** and **peace** every time I return*); Satisfaction was realized via expressions like *satisfied, concern, interest, proud*, etc. (e.g., *Most visitors expressed **satisfaction** with delicious local food . . .*). This indicates that students have learned how to express their personal emotions more clearly and richly.

More remarkably, students have known how to make their writing more interesting and attractive by employing more Appreciation resources (evaluation of things and entities). Appreciation realizations in the post-test writing were much more varied than those in the pre-test. For instance, to realize Reaction, such adjectives as *pleasing, vibrant, attractive, impressive, surprising*, etc. were frequently employed (e.g., *The rice paddies . . . enable the village to have an **impressive** background that is actually very **pleasing** to the eye of the residents and in effect provides an exceptionally **tranquil** setting for its inhabitants to live their lives to the fullest; . . . local markets are filled with **vibrant** colors, showcasing traditional crafts and **delicious** food*); realizations of Composition included *balanced, simple, dynamic, logical, natural*, etc. (e.g., *The town has a **dynamic** and **well-balanced** appearance thanks to its pleasant mixture of modern and old buildings; Its rapidly growing economy and modernization are beautifully **balanced** by the preservation of tradition, offering both tremendous progress and cultural richness . . .*) and realizations of Valuation were really rich with *crucial, significant, valuable, good, rich, effective*, etc. (e.g., *Hanoi is a city with a **rich** historical background . . . ; . . . the food here is prepared with fresh ingredients and has a **unique** flavor that reflects the local traditions; Hanoi has a **rich** history, with many **significant** sites that still stand today, such as the Thang Long Imperial Citadel*). Thus, an evaluation of a student in the pre-test (e.g., *Thai Nguyen has **large** tea fields . . . , people are **friendly** . . .*) turned into a livelier and more attractive one in the post-test (e.g., ***Friendly** locals, **bustling** markets and the **serene** tranquility of tea fields make Thai Nguyen a **truly special** place, where tradition and progress exist in **perfect** harmony.*)

In short, after 12-week intervention with the introduction and practice of the Appraisal framework, the quality of students' writing in terms of Attitudinal expressions improved dramatically, making their writing more interesting and convincing.

(b). *Effects on Graduation Resources*

Beside using realizations to express Attitude, students also learned how to clarify their views by upscaling or downscaling their evaluation with Graduation resources. A MANOVA test, therefore, was also carried out to check how EG and CG students performed their understanding and usage of Graduation in the post-test. According to reports (Table 5), there was a statistically significant overall difference between the EG and CG groups on the combined dependent variables, *Wilks' Λ* = .588, $F(2,57) = 19.975$, $p < .001$. Follow-up tests on Graduation's categories showed similar results for Force ($F(1,58) = 24.916$, $p < .001$, partial $\eta^2 = .300$), Focus ($F(1,58) = 20.754$, $p < .001$, partial $\eta^2 = .264$) and all Graduation resources ($F(1,58) = 39.102$, $p < .001$, partial $\eta^2 = .403$).

TABLE 5
MANOVA RESULTS FOR GRADUATION FREQUENCIES (PER 100 WORDS) AT POST-TEST

Category	Group	Mean $\pm$ SD	$F(1,58)$	p -value	Partial η^2
Overall Graduation	EG	6.4450 $\pm$ 1.49973	39.102	<.001	.403
	CG	3.9820 $\pm$ 1.55085			
Force	EG	4.6780 $\pm$ 1.41445	24.916	<.001	.300
	CG	3.0420 $\pm$ 1.10540			
Focus	EG	1.7670 $\pm$.72611	20.754	<.001	.264
	CG	.9400 $\pm$.67927			

These results suggest that the EG group employed significantly more Graduation resources overall than the CG in the post-test. This difference was evident in both dimensions of Force and Focus.

Deeper qualitative analysis of EG group's writing provides more detailed of how EG students' employment of Graduation resources has changed after the intervention. As can be seen from Table 6, the frequency of Graduation resources reports a dramatic increase by 3.37 percentage points from 3.07% to 6.44%. This increase is observed in both Force and Focus categories, whose frequencies have approximately doubled in the students' post-test.

TABLE 6
GRADUATION FREQUENCIES IN PRE- AND POST-TESTS AMONG EG STUDENTS (PER 100 WORDS)

EG group	Graduation resources		Total
	Force	Focus	
Pre-test	2.31	0.76	3.07
Post-test	4.68	1.77	6.44

This is evidence that students of EG group have acquired knowledge and skills to intensify their evaluations through diverse Graduation realizations. For Force, students could use both Intensifiers (amplify the degree of an emotion, quality, or process) and Quantifiers (amplify quantity) to describe their feelings and judgements in a more vivid and colorful way. Various realizations of Intensifiers and Quantifiers were identified from the post-test essays such as *greatly*, *extremely*, *very*, *immensely*, etc. (Intensifiers) and *various*, *considerable*, *many*, *substantial*, *rich*, etc. (Quantifiers). As such, after the intervention, an EG student would write "*The temples and historical sites are **immensely** fascinating*", "*The natural surroundings are **strikingly** beautiful with an **extremely** aesthetic bay and **many** gorgeous beaches . . .*", or "*It is a city with a **really** rich history and is **rather** well-known for its production of green bean cakes*" instead of "*The city is beautiful*" in the pre-test.

Besides, students' use of Focus has also been improved. To sharpen or soften evaluation of unscalable entities, students knew how to use such words as *actually*, *genuine*, *pure*, *truly*, *kind of*, *somewhat*, etc. For example: "*Thai Nguyen City is a **truly** unforgettable destination . . .*"; "*. . . the town is **somewhat** isolated . . .*"

Thus, proper employment of Graduation resources helps students express their Attitude more persuasively and effectively. The intervention did have positive impact on EG students' cognitive understanding and practical use of evaluative lexical resources.

(c). *Effects on Overall Writing Quality*

Concerning the overall writing quality, ANCOVA results (Table 7) revealed a statistically significant difference in overall writing scores on the post-test between the groups, after controlling for pre-test scores ($F(1,57) = 12.292$, $p < .005$, partial $\eta^2 = .177$). The EG achieved significantly higher mean scores than the CG, indicating a clear improvement in general writing quality.

TABLE 7
ANCOVA RESULTS FOR POST-TEST SCORES

Dependent variable	Covariate	Adjusted mean (EG)	Adjusted mean (CG)	$F(1, 57)$	p -value	Partial η^2
Scores (Post-test)	Scores (Pre-test)	7.4667	7.0043	12.292	.001	.177

Students' writings were assessed according to an analytic rubric including specific criteria (Content, Organization, Grammar, and Vocabulary) rated on a 1–10 scale (Appendix B). These criteria were selected to ensure the objectivity and consistency of writing evaluation for both the experimental and control groups. To explore the potential relationships between the use of evaluative language and writing performance, Pearson correlation analyses were

conducted. In particular, two separate correlation analyses were carried out to examine the relationships between frequencies of Attitude and Graduation and the post-test writing scores of the experimental group (Table 8).

TABLE 8
PEARSON CORRELATION BETWEEN FREQUENCIES OF ATTITUDE AND GRADUATION AND POST-TEST SCORES

Variable pair	<i>r</i>	Sig. (2-tailed)	<i>n</i>
Attitude – Scores	.900	<.001	30
Graduation – Scores	.884	<.001	30

As shown in Table 8, there was a strong, positive correlation between the frequency of Attitude resources and post-test writing performance ($r = .90, p < .001$), suggesting that students' writings with more Attitudinal expressions tended to achieve higher post-test scores. Similarly, the frequency of Graduation resources in the writing was strongly correlated with its score ($r = .88, p < .001$), indicating that students with higher overall writing scores also performed better employment of Graduation resources to clarify their personal stance.

It can be concluded that students' writing quality has been much improved with the application of the Appraisal framework. Students know how to express their stance and make their writing impressive with the employment of various evaluative resources.

D. Student Perceptions

A questionnaire consisting of 12 items (with 10 closed-ended and 2 open-ended questions) was administered to assess students' perception of the application of the Appraisal framework in their writing course. The questionnaire was only collected from the EG group ($n = 30$).

Data from the questionnaire (Table 9) indicated generally positive responses towards learning and applying the Appraisal framework in writing lessons. One hundred percent of students agreed or strongly agreed that after studying the Appraisal framework, their understanding of ways to express attitude and stance became very good ($M = 4.46$); they could identify lexical resources to express emotions more clearly ($M = 4.30$); and their writing became livelier and more convincing ($M = 4.23$). Most students agreed or strongly agreed that the Appraisal framework helped them to improve the performance of evaluative vocabulary in their writing ($M = 4.10$) and they became more confident when writing English essays ($M = 4.40$). Meanwhile, the impact of the Appraisal framework in building more logical and coherent essays received a moderately high score ($M = 3.70$), indicating that while students were generally confident in their stance expression and lexical improvement, they did not highly appreciate the effect of the Appraisal framework in organizing their writing.

TABLE 9
DESCRIPTIVE STATISTICS FOR QUESTIONNAIRE ITEMS

Items	<i>n</i>	Minimum	Maximum	Mean	Std. Deviation
Understanding of evaluative expressions in English writing	30	4.00	5.00	4.4667	.50742
Appraisal framework's use in realizing emotional expressions in English writing	30	4.00	5.00	4.3000	.46609
Improvement of evaluative vocabulary usage	30	3.00	5.00	4.1000	.71197
Better and more convincing writing	30	4.00	5.00	4.2333	.43018
More confident in writing English	30	3.00	5.00	4.4000	.62146
More logic and coherent writing	30	3.00	5.00	3.7000	.59596
Difficulty in memorizing and using evaluative expressions	30	1.00	4.00	2.3333	1.12444
Duration of the course	30	1.00	3.00	2.1333	.73030
Overall evaluation on the effectiveness of the course	30	4.00	5.00	4.4333	.50401
Valid <i>N</i> (listwise)	30				

However, a small number of students also acknowledged initial difficulties in memorizing and differentiating the terminology (Table 10). More specifically, some students had difficulty understanding concepts in the Attitude system (16.67%) and Graduation system (20%). Few students (10%) admitted that they found it challenging to apply these evaluative items in their writing and that they need more examples to illustrate. Nearly a third of students (26.27%) reported difficulty practicing the framework and responding to teachers' tasks.

TABLE 10
DIFFICULTIES IN LEARNING AND APPLYING THE APPRAISAL FRAMEWORK

Items	Yes (<i>n</i>)	No (<i>n</i>)	Total	Yes (%)	No (%)
Understanding concepts of Attitude	5	25	30	16.7	83.3
Understanding concepts of Graduation	6	24	30	20.0	80.0
Applying the Appraisal framework	3	27	30	10.0	90.0
Practicing and responding	8	22	30	26.7	73.3
Lack of examples	3	27	30	10.0	90.0

As a result, about 30% of students (Table 9) expressed they needed more time to study the framework as the duration for the framework in the course was not enough. Some students suggested increasing the practice time and providing more specific illustrative examples.

Overall, one hundred percent of students assessed that the framework's impact on English writing skills was Good or Very good ($M = 4.43$), partly indicating the achievement of the intervention period (Table 9).

Qualitative analysis of open-ended responses highlighted several key themes. Being asked to share personal feelings or experience in the process of studying and applying the framework, students tend to have positive feedback. One student shared: "Before, I just wrote what I thought. Now, I think about how to say it . . . Using words like *"significant"* or *"slightly"* makes my point stronger or softer. Appraisal gave me the words for these choices." Another student wrote: "The Appraisal framework helps me better understand how to express emotions and opinions in writing".

Secondly, students recognized the practical application of the framework in constructing arguments, particularly using Attitude for evaluation and guiding the reader. One student noted: "Learning about Attitude, especially Judgement and Appreciation, helped me evaluate the topic better, not just give my feeling." Another student wrote: "I feel my writing is more convincing after learning about the elements of Appraisal".

Thirdly, challenges and the learning curve were also mentioned, mainly related to the initial complexity of the terminology. One student stated: "The names were confusing at first. Affect, Judgement, Appreciation ... and all the small types. It took time. Sometimes in the essay, I wasn't sure if it was Judgement or Appreciation, but the main idea of evaluating was helpful anyway." Some students suggested that there should be more time and practice exercises to help them better memorize and apply the framework.

Overall, despite initial challenges, the majority of EG students found learning the Appraisal framework a valuable and beneficial experience for improving their English writing skills.

V. DISCUSSION

Results of this study provide clear empirical evidence for positive effects of integrating the Appraisal framework into English writing instruction to improve writing skills for Vietnamese university students in the studied context. Both quantitative and qualitative analyses consistently showed that the group of students exposed to explicit instruction on Attitude and Graduation resources demonstrated significant advantages over the control group on several important dimensions after the intervention. Specifically, EG students employed Attitude resources, particularly Affect and Appreciation, more frequently and appropriately, enabling them to make more vivid and persuasive arguments. They also show a more varied and sophisticated use of Graduation resources to modulate the intensity and range of utterances, demonstrating greater control over nuances of meaning. These improvements in the use of Appraisal resources simultaneously led to significantly higher overall writing quality for the EG. In addition, positive feedback from EG students, despite acknowledging initial difficulties, confirmed the cognitive value and practical applicability of learning the Appraisal framework.

Results of this study are consistent with previous studies on the effectiveness of the Appraisal framework in teaching English writing skills (Derewianka, 2007; Hood, 2010; Hu & Choo, 2015; Liu, 2010). EG students' greater use of Appreciation resources appropriately supports observations of the role of these types of Attitudes in constructing arguments and demonstrating critical thinking (Hood, 2010). The significant increase in the use of Graduation shows that students not only learn how to express Attitude but also know how to adjust the level of expression in writing, helping to improve the quality of writing comprehensively. Hu and Choo's (2015) study in Singapore also noted similar improvements in students' writing skills when applying the Appraisal framework, especially in the use of Graduation categories to enhance persuasiveness. Similarly, Liu (2010) found that the Appraisal framework helped Chinese students improve their ability to express emotions and opinions in their essays. Thus, the increased use of evaluation elements makes writing more vivid, convincing, and personalized. This study adds further evidence of the pedagogical application of the Appraisal framework in the specific EFL context of Vietnam, where the development of evaluative and interactive expressive competence in written English remains a challenge.

The effectiveness of the Appraisal framework-based intervention can be attributed to several factors. First, providing students with a framework to discuss language choices related to evaluation appears to have increased their awareness and conscious control over these aspects of their writing. Instead of focusing solely on structure and grammar, students were guided to think about the meaning and effectiveness of word choices in expressing their point of view and connecting with the reader. Second, analyzing sample texts through the lens of the Appraisal framework provided concrete examples and models of how experienced writers effectively use these resources. Third, deliberate practice activities focusing on the application of the concepts of Attitude and Graduation provided opportunities for students to experiment with and reinforce the use of these resources in specific writing contexts.

From the above results and interpretations, some important pedagogical implications can be drawn for the teaching of English writing at tertiary level in Vietnam.

First, considering systematically integrating elements of the Appraisal framework into EFL writing curricula is a potential direction. This does not necessarily mean replacing existing teaching methods but rather adds an important dimension that focuses on interpersonal meaning.

Second, specific teaching activities might include using the Appraisal framework as a tool for analyzing sample texts, designing small exercises that help students practice identifying and producing different Attitude and Graduation expressions, and incorporating Appraisal-related criteria into teacher and peer feedback.

Third, it is also necessary to develop teaching materials that are appropriate to the Vietnamese context, which may use simplified terminology or focus on the core concepts initially.

Last but not least, professional development programs are needed to equip EFL teachers with the knowledge and necessary skills to understand and effectively apply the Appraisal framework in the classroom, overcoming barriers of initial familiarity or apprehension towards SFL-based approaches.

VI. CONCLUSION

Through explicit pedagogical focus on the language of evaluation - especially the systems of Attitude and Graduation - the study demonstrated statistically significant improvements among experimental group students in their ability to express assessments appropriately, producing higher-quality essays compared to the control group receiving traditional instruction. Notwithstanding the observed difficulties in learning, most students thought the framework was a useful resource that improved their metalinguistic awareness.

It is recommended that educators integrate the Appraisal framework into their English writing curriculum systematically, combined with varied practice activities and detailed feedback to optimize learning outcomes. Teachers need to be properly trained on the Appraisal framework to guide students more effectively.

However, this study also has some limitations that should be acknowledged. The quasi-experimental design did not allow for complete control of potential confounding factors due to the lack of random assignment, although ANCOVA was used to control initial differences. The study sample was limited to students at one university, so the generalizability of the results to other student populations in Vietnam should be considered with caution. The intervention period (12 weeks) may not have been long enough to observe long-term effects or complete mastery of the Appraisal resources.

Based on the results and limitations of the study, some directions for further research can be suggested. Longitudinal studies that track the development of students' Appraisal competencies over a longer period of time would provide further insight. Replication of the study in universities and with groups of students with different English proficiency levels in Vietnam would help verify and expand the generalizability of the results. Exploring the application of the Appraisal framework in teaching other genres (e.g., research reports, literature reviews) is also a promising direction. Future research could also delve into teachers' perspectives and challenges when implementing this method in teaching practice in Vietnam.

APPENDIX A. WRITING SYLLABUS INTEGRATING APPRAISAL FRAMEWORK INSTRUCTION

TABLE A
WRITING SYLLABUS INTEGRATING APPRAISAL FRAMEWORK INSTRUCTION

Week	General writing	Appraisal framework
1.	Self-introduction paragraph; Topic sentences & supporting details;	
2.	Descriptive paragraph; Organize descriptive writing	Introduce the Appraisal framework; Attitude – Affect (emotion words)
3.	Short biography; Structure biography	Attitude – Judgement & Appreciation
4.	Prediction essay; Organize logically	Graduation – Force
5.	Report on endangered species	Attitude + Graduation
6.	Narrative; Sequencing words	Affect (emotions in storytelling)
7.	Formal & informal emails; Conventions of email writing	Force (adjusting tone intensity)
8.	Report on spending habits	Focus (precision vs. vagueness)
9.	Cause–effect essay; Structure essay	Attitude + Graduation
10.	Argumentative essay	Attitude + Graduation
11.	Formal letter/request	Attitude + Graduation
12.	Final essay	Apply Attitude + Graduation

APPENDIX B. RUBRIC FOR ESSAY ASSESSMENT

TABLE B
RUBRIC FOR ESSAY ASSESSMENT

Score	Content	Organization	Vocabulary	Grammar
10	Task fully addressed with depth and nuance	Clear organization; wide range of cohesive devices	Wide vocabulary; effective collocations	Very wide range; complex structures used flexibly
9	Rich ideas, well-developed	Logical organization; clear paragraphing; varied linking words	Very good range; mostly precise and appropriate	Very few minor errors; good mix of sentence types
8	Task fully covered; ideas clear, logical	Well-structured; clear paragraphing; cohesive devices mostly appropriate	Good range; limited collocations	Generally accurate; occasional errors do not affect meaning
7	Task addressed with relevant ideas	Adequate structure; limited range of cohesive devices	Adequate vocabulary; some repetition	Attempts at complex sentences, with minor mistakes
6	Cover task with basic ideas	Simple but clear structure; cohesive devices simple	Limited range; noticeable repetition	Limited range; frequent errors in word forms; tense control inconsistent
5	Partial task coverage; ideas vague, little support	Paragraphs poorly separated; few cohesive devices	Restricted vocabulary; frequent errors in word choice	Simple repeated structures; complex structures often incorrect
4	Very limited ideas; mostly underdeveloped	Unclear organization; little paragraph control; minimal linking words	Very limited vocabulary; frequent errors	Very limited control; frequent basic errors; mostly simple sentences
3	Few ideas; hardly address task	Very poor structure; no clear paragraphing	Extremely limited range; many incorrect or inappropriate words	Very poor control of tenses; sentences often incomplete or incorrect
2	Task largely ignored	Almost no structure; no linking words	Almost no usable vocabulary	Rare complete sentences
1	Fail to address task	No organization; incoherent	Irrelevant words	No control of grammar

REFERENCES

- [1] Alhuthali, M. (2024). Analysis of two EFL writing genres using appraisal theory. *British Journal of English Linguistics*, 12(3), 15–23. <https://doi.org/10.37745/bjel.2013/vol12n31523>
- [2] Coffin, C. (2003). Reconstruals of the past – Settlement or invasion? The role of Judgement analysis. In J. R. Martin & R. Wodak (Eds.), *Re/reading the past: Critical and functional perspectives on time and value* (pp. 219–246). John Benjamins.
- [3] Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. Continuum.
- [4] Coffin, C., & O'Halloran, K. A. (2005). Finding the global groove: Theorizing and analyzing dynamic reader positioning using appraisal, corpus and a concordancer. *Critical Discourse Studies*, 2(2), 143–163. <https://doi.org/10.1080/17405900500283607>
- [5] Derewianka, B. (2007). Using appraisal theory to track interpersonal development in adolescent academic writing. In A. McCabe, M. O'Donnell, & R. Whittaker (Eds.), *Advances in language and education* (pp. 142–165). Continuum.
- [6] Duy, T. V., & Trang, B. T. T. (2022). Investigating difficulties of English-major students in writing academic essays. *European Journal of Foreign Language Teaching*, 6(3), 38–50. <https://doi.org/10.46827/ejfl.v6i3.4512>
- [7] Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- [8] Hood, S. (2004). Appraising research: Evaluation in academic writing. In M. Thompson & L. Alba-Juez (Eds.), *Evaluation in context* (pp. 67–92). John Benjamins Publishing Company.
- [9] Hood, S. (2010). *Appraising research: Evaluation in academic writing*. Palgrave Macmillan.
- [10] Hu, G., & Choo, L. (2015). Applying the Appraisal framework in teaching English writing: A pedagogical perspective. *Journal of Language Teaching and Research*, 6(3), 512–520. <https://doi.org/10.17507/jltr.0603.01>
- [11] Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192. <https://doi.org/10.1177/1461445605050365>
- [12] Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- [13] Lee, S. H. (2015). Evaluative stances in persuasive essays by undergraduate students: Focusing on APPRECIATION resources. *Text & Talk*, 35(1), 49–76. <https://doi.org/10.1515/text-2014-0029>
- [14] Liu, J. (2010). The application of Appraisal theory in English writing teaching. *English Language Teaching*, 3(1), 123–128. <https://doi.org/10.5539/elt.v3n1p123>
- [15] Liu, Y. (2013). Evaluation in Chinese university EFL students' English argumentative writing: An appraisal study. *Journal of English for Academic Purposes*, 12(4), 335–347. <https://doi.org/10.1016/j.jeap.2013.09.003>
- [16] Macken-Horarik, M. (2003). Appraisal and the special instructiveness of narrative. *Text & Talk*, 23(2), 285–312. <https://doi.org/10.1515/text.2003.012>
- [17] Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- [18] Nghi, N. P. A., & Truong, T. N. (2023). Difficulties of Vietnamese students in learning academic writing. *International Journal of Asian Education*, 4(2), 106–114. <https://doi.org/10.46966/ijae.v4i2.340>
- [19] Nguyen, B. H. (2021). Evaluative language in conclusion sections of Vietnamese linguistic research articles. *VNU Journal of Foreign Studies*, 37(3), 111–122. <https://doi.org/10.25073/2525-2445/vnufs.4643>

- [20] Nguyễn, T. K. N., & Nguyễn, T. H. L. (2020). A preliminary study on attitude in English and Vietnamese media texts in the light of appraisal theory. *Journal of Foreign Studies*, 36(3), 101–117. <https://js.vnu.edu.vn/FS/article/view/4559>
- [21] Phuong, W. T. N. (2021). Difficulties in studying writing of English-majored sophomores at a university in Vietnam. *European Journal of Education Studies*, 8(10), 1–14. <http://doi.org/10.46827/ejes.v8i10.3962>
- [22] Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum Associates.
- [23] Swain, E. A. (2010). Getting engaged: Dialogistic positioning in novice academic discussion writing. In E. A. Swain (Ed.), *Thresholds and potentialities of systemic functional linguistics: Multilingual, multimodal and other specialized discourse* (pp. 291–317). Trieste, Italy: EUT Edizioni Università di Trieste.
- [24] Tran, X. H. (2023). Investigating difficulties faced by second-year English major students in learning English writing skills at Hanoi university of natural resources and environment. *Journal of Educational Equipment: Applied Research*, 1(298), 75–82. <https://vjol.info.vn/index.php/tctbgd/article/view/89754>
- [25] Wu, S. M., & Allison, D. (2007). The use of engagement resources in high- and low-rated undergraduate geography essays. *Journal of English for Academic Purposes*, 6(3), 254–271. <https://doi.org/10.1016/j.jeap.2007.09.006>
- [26] Yang, Y. (2016). Appraisal resources in Chinese college students' English argumentative writing. *Journal of Language Teaching and Research*, 7(5), 1002–1013. <https://doi.org/10.17507/jltr.0705.23>



Hong B. Nguyen was born in Hanoi (Vietnam) on 7th August 1981. She earned the Bachelor's degree in English language teaching and Master's degree in English from University of Language and International Studies in 2003 and 2009 respectively. She got a degree in Business Administration from Thuongmai University in 2015. She earned her PhD in Contrastive Linguistics in 2023 from Academy of Social Science, Vietnam.

She has been working as a teacher of English at Thuongmai University for 20 years. She has published five research articles in Scopus indexed journals and 16 ones in journals and proceedings of Vietnamese and international conferences, 3 books and 4 research projects at university level. Her main research interests are systemic functional linguistics (especially evaluative language), contrastive analysis and pedagogy. Email: hongnguyen.dhtm@tmu.edu.vn ORCID iD: <https://orcid.org/0009-0006-6440-898X>

The Impact of the Flipped Classroom Model on Developing EFL Students' Four Language Skills: An Experimental Study

Rasha Dorgham

Department of English Language and Literature, College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

Lamya Bin Obiad

College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

Abstract—This study examined the effects of flipped instruction on the reading, writing, listening, and speaking skills of English as a Foreign Language (EFL) learners. The educational environment continues to evolve and adapt under exceptional and timely situations to enhance language acquisition, with innovative teaching approaches like the flipped classroom development that promote student autonomy and involvement. In recent years, learning outcomes in students have indicated that flipped learning—or the presentation of instructional content outside the classroom so that classroom time can be used for practice, participation, and interaction—has significantly influenced language learning. At Fatima Al-Zahraa Secondary School, a quasi-experimental research design was adopted for this study. The participants were female high school juniors, divided into two groups—an experimental group and a control group. The results indicate that the use of flipped classroom pedagogies in EFL settings increases students' engagement and enables the all-round development of language skills. The study also includes implications for the practice of the flipped instruction model and proposals for further investigation into extended flipped instruction, especially in relation to language proficiency levels and diverse teaching contexts.

Index Terms—flipped classroom, EFL learners, student engagement, language education, instructional innovation

I. INTRODUCTION

English as a Foreign Language (EFL) instruction has become a significant part of their education systems. However, long-standing issues with teaching EFL programs that have significant effects on student learning results, engagement, and overall language skill development continue to plague academic institutions worldwide. Students struggle to develop their integrated language skills and exhibit low levels of active participation, interaction, and self-directed learning in class activities. Moreover, investigations into traditional educational practices, including teacher-centered instruction, passive reception of knowledge, and a lack of authentic language practice, are all contributing factors to students' disengagement and leading them to fail at achieving successful learning outcomes (Chen Hsieh et al., 2017; Öztürk & Çakıroğlu, 2021). On the other hand, the COVID-19 pandemic exacerbated existing challenges with student engagement when educational institutions used remote or mixed learning models that made traditional approaches to instruction in the classroom problematic concerning student motivation and language development (Oraif & Elyas, 2021). As a result, investigations into whether the flipped classroom is a potential solution to this problem have greatly increased.

The flipped classroom approach is defined by the reverse of the usual sequence for teaching and learning. In this pedagogic approach, the learner reads instruction materials, for example, video lecture readings or multimedia resources, before receiving face-to-face class time and then spending time in classes to do active learning (for example, group work/collaborative work, problem-solving, or instructor-led support) (Bergmann & Sams, 2012; Lo & Hew, 2017). Based on that, the current research is seeking to answer the question, "How much does the flipped classroom approach help to develop the four language skills (reading, writing, listening, and speaking) and facilitate student engagement in the EFL classroom?" To answer this question, the study utilized both pre- and posttests and a questionnaire to investigate whether or not the flipped learning approach can really improve the teaching and learning of English as a Foreign Language in secondary schools.

II. LITERATURE REVIEW

A. The Flipped Classroom Model in Education

The underlying pedagogy for flipped learning is embedded within the frameworks of constructivist learning theories that emphasize active construction of knowledge, social interaction, and learner autonomy. Lo and Hew (2017) note that

the flipped model incorporates elements of Bloom's Taxonomy by creating students who are expected to engage with lower-order cognitive tasks of remembering and understanding outside of class and reserving their face-to-face time for learning that engages higher-order thinking skills of applying, analyzing, evaluating, and creating. Furthermore, the nature of this pedagogical approach encourages more meaningful learning and aids cognitive development (Evseeva & Solozhenko, 2015).

Literature on the advantages of flipped instruction has grown in a variety of educational disciplines. Studies show evidence of flipped classrooms resulting in increased engagement, better performance, improved satisfaction with learning, and the development of self-regulated learning skills (Akçayır & Akçayır, 2017). The model provides students with flexibility to learn at their own speed, review content as necessary, and come to class prepared for learning.

However, as with any innovative curriculum, there are relevant challenges educators need to consider in implementing the flipped classroom model. Lo and Hew (2017) previously identified challenges that included inconsistent student acceptance of a non-traditional learning experience, a lack of student preparation before class, a lack of technology resources to learn with, an increased preparation workload for teachers in being prepared with content, and the assurance of student accountability for learning before class (Akçayır & Akçayır, 2017). To implement the flipped classroom model, educators must consider pre-service planning, appropriate planning of the communication of classroom expectations, availability of appropriate technology infrastructure, and ongoing support for staff and students. Additionally, learning modules must be responsive to student needs and interests while meeting an acceptable learning and performance outcome. In-class activities should also maximize the effectiveness of students' face-to-face learning experience.

B. The Flipped Classroom in Language Learning

Flipped classroom approaches in language education have received a great deal of attention due to their application to concrete problems experienced in foreign language learning. For example, Öztürk and Çakiroğlu (2021) conducted a study analyzing the design and implementation of flipped learning incorporating self-regulated learning strategies within an English as a Foreign Language (EFL) context. Their research findings reported that the flipped model effectively enhanced students' metacognitive awareness, strategic learning behaviors, and autonomous learning competencies.

Fathi and Rahimi (2022) investigated the influence of flipped classroom teaching on the complexity, accuracy, and fluency of EFL students' writing. Results of the study indicated that flipped instruction had a statistically significant effect on the three domains of writing performance. In most cases, students who learned through flipped classroom teaching produced more syntactically complex and fluent texts with fewer grammatical mistakes compared to students in the control sample. The authors indicated that the flipped model yielded positive opportunities for improving writing proficiency through the out-of-class preparation and the use of the out-of-class work to practice in the classroom.

This is timely, as Qiu and Luo (2022) studied the effects of flipped listening instruction on the listening performance and listening anxiety of Chinese EFL college students. The researchers found that students in the flipped classroom environment had significantly higher scores in listening comprehension and less listening anxiety than their traditional classroom counterparts. The researchers mentioned that the flipped classroom design encouraged the students to preview their listening, develop listening strategies, and receive teachers' support to complete tasks that they would have had difficulty completing on their own, hence increasing performance and decreasing anxiety.

C. The Relationship Between the Flipped Classroom and Academic Achievement

Ample research shows positive relationships between flipped classrooms and academic achievement in a number of educational contexts and subjects (for example, Chen Hsieh et al., 2017). Eltahir and Alsalhi (2025) confirm the positive effect of flipped classroom on students' achievement, motivation, and classroom engagement in higher education. Teng (2018) further reported that EFL students displayed notably enhanced speaking ability following a flipped-classroom experience. Pre-test and post-test comparisons demonstrated significant improvements in speaking skills, specifically pronunciation, grammar, vocabulary, discourse, and interaction. This study showed that increased use of class time for speaking practice and having students' complete lectures and readings led to increased speaking ability.

Qiu and Luo (2022) investigated the listening comprehension outcomes of flipped and non-flipped classrooms. Notably, the students in the flipped class attained significantly higher scores on a standard listening test/scale and made gains in the composites of main ideas, details, inference, and pragmatics. The authors noted that the flipped method addressed the issues of traditional listening teaching by providing sufficient time for preparation and enabling learners to practice in a more focused way during the lesson.

In general, the studies to date indicate positive effects of the flipped classroom model on students' learning performance. Although the positive characteristics of the flipped classroom may be obvious, it has been argued that its effect on academic results could be influenced by other variables, such as quality of assignments before class, design of activities during class, attributes of students, technological infrastructure, and level of teachers' preparedness for the methodology of flipped learning (Akçayır & Akçayır, 2017).

D. The Relationship Between the Flipped Classroom and Student Engagement

Engagement is a key determinant in learning and success. Engagement in learning has three dimensions: behavioral engagement (students' involvement in learning activities), emotional engagement (students' feelings and emotional responses to learning), and cognitive engagement (students' cognitive efforts during learning) (Mok & Flynn, 2019).

Flipped classroom pedagogy is also anticipated to be especially effective in fostering engagement in all three aspects (Khaleel et al., 2020). Also, Mizz et al. (2025) highlights the effectiveness of the flipped classroom in enhancing student engagement while strengthening inclusive learning environments.

Chen Hsieh et al. (2017) analyzed the motivation and engagement of learners in both EFL inverted classrooms and in traditional classrooms. The surveys and qualitative interview results demonstrated that the level of motivation, interest in learning, and satisfaction with classroom experiences in flipped classes were much higher. The participants liked the fact that they received the opportunity to practice interactively during the session, which helped them continue to stay involved and have a positive attitude towards learning (Ardi & Rianita, 2022). Moreover, the flipped classroom method promotes active involvement, student independence, interaction, and reflection on learning (Alsalam, 2025).

E. Related Previous Studies

Numerous studies have investigated the utilization of flipped classroom methodologies in EFL contexts, building on what has been learned about implementation, learning outcomes, and contextual factors that influence efficacy.

Öztürk and Çakiroğlu (2021) examined the correlation between flipped learning and self-regulated learning in an EFL class. They observed that learners who took part in flipped learning, together with the self-regulated learning method, achieved better language skill results and were more active in learning. The experimental group participants showed greater use of their metacognitive thinking and study strategies than the control group and possessed more self-control over their learning. They concluded that flipped learning and SLL are two effective teaching strategies that can be integrated to enhance and deepen language learning.

In their study, Fathi and Rahimi (2022) concentrated entirely on the applicability of the flipped class method in augmenting writing proficiency. Based on a quasi-experimental pretest/posttest control group design, they adopted a stringent approach to investigating the impact of flipped teaching on writing complexity, accuracy, and fluency. They also examined core linguistic areas such as syntactic complexity, error distribution, and fluency measures to offer strong evidence of developing writing within the flipped classroom model. The authors highlighted that both in-class and out-of-class, flip-based instruction with teacher feedback facilitated students' writing development by providing more opportunities to engage in extended practice and sustained written performance.

Similarly, Qiu and Luo (2022) examined the effects of the flipped classroom on listening anxiety and listening comprehension. Results demonstrated that the flipped pedagogy could enhance the listening task performance and decrease task-based listening anxiety in an EFL environment. The results were explained by the instructional design of flipped lessons, as they alleviated the cognitive load involved in acquiring new content and language use, enabled learners to plan and practice outside of the class, and provided learner support within the classroom. To date, the research evidence pooled from these earlier studies serves as a compelling rationale for offering flipped classroom approaches to EFL.

III. METHODOLOGY

A. Research Design

This study used a quasi-experimental research design employing both pre-test and post-test data collection involving two groups. This method is appropriate for measuring the effectiveness of a pedagogical intervention (i.e., the flipped classroom model) on dependent variables for the four language skills (reading, writing, listening, and speaking) and student engagement in EFL contexts.

The experimental and control groups both received instruction in the four language skills; however, the difference was in the teaching approach.

The study's participants were female students in their third year at Fatima Al-Zahraa Secondary School, a public educational institution in Alexandria, Egypt, during the academic year 2024-2025. The study included 60 participants, who were divided equally into two groups: an experimental group and a control group, each with 30 participants.

Furthermore, the female participants were aged 16 to 17 years; however, three students were 13 years old and represented typical demographic characteristics of third-year secondary school students in Egyptian school settings. Questionnaire data obtained at the pre-study administration provided evidence that 78.3% of the participants (47 students) indicated that they make regular use of educational technology platforms, while 21.7% of the participants (13 students) reported very limited experience with technology in the learning environment. However, all students had basic levels of technological literacy skills necessary for engagement with digital learning materials as required.

Purposive sampling procedures were used to identify participants who had similar academic performance levels and backgrounds for English language proficiency.

Additionally, an independent samples t-test statistical analysis confirmed there were no statistically significant differences ($p > 0.05$) between experimental and control groups. Both groups are homogeneous. This study involved the use of three main instruments:

(a). Language Skills Assessment (Pre-Test and Post-Test)

The research team developed a test to measure the English language skills (reading, writing, listening, and speaking) of the participants, with each of the four sections worth 25 points. Ten multiple-choice questions assessed reading comprehension, with five multiple-choice questions assessing inferential comprehension and five short-answer questions

assessing some level of analysis of text.

The listening comprehension section was designed to test how students comprehend audio material. It included 10 multiple-choice questions assessing listening comprehension with audio recordings—five true/false questions identifying details and five short answer questions requiring participants to draw inferences from the listening passage.

In the speaking part of the test, students were asked to describe a picture, participate in a role-play conversation, and express an opinion. Each task was assessed via a set of rubrics that evaluated fluency, accuracy, pronunciation, vocabulary range, and communication.

The tests were administered both before and after the instruction intervention to evaluate any change in language proficiency. Content validity was established by experienced EFL instructors who served as experts. A reliability analysis yielded a Cronbach's alpha coefficient of 0.87, indicating strong internal consistency.

(b). Student Engagement Questionnaire

The research utilized a modified version of the student engagement scale created by Oraif and Elyas (2021) with 25 items.

The first component was Behavior Engagement, which included nine questions about how often or how much the students participated in class, how often they finished their assignments, and how actively they participated in learning activities. The second component was Emotional Engagement, which contained eight items that measured students' interest in learning English, their enjoyment of learning activities, and their attitudes toward studying English. Finally, the third component was Cognitive Engagement, using eight items that measured students' use of deep strategies, their metacognition, and their hard work on difficult tasks.

Participant responses were rated on a five-point Likert scale, beginning at 1 (Strongly Disagree) and ending at 5 (Strongly Agree). The total score could range from 25 to 125 points for all 25 items. The questionnaire was given to both groups after the instructional intervention was completed. The instrument had a previous reliability coefficient of $\alpha = 0.92$. Pilot testing with the current sample produced a reliability coefficient of $\alpha = 0.90$.

Additionally, to explore students' perceptions and attitudes towards their learning experiences, a 10-item adjunct satisfaction instrument (with items measuring aspects of instructional effectiveness, quality of learning materials, opportunities for practice, teacher support, and satisfaction with the general learning experience) supplemented qualitative perspectives on the students' subjective experiences.

B. Procedures

The implementation of the research study took place over a period of seven weeks during the second semester of the 2024-2025 academic year and followed a thoughtfully planned sequence:

Week 1: Training Stage for Flipped Classroom Implementation

Before the intervention, the experimental group went through a one-week training period to help students become acquainted with the flipped classroom approach. As part of this training stage, students attended an orientation meeting about the flipped classroom model, which described the approach's structure and expectations. Students were also introduced to the LMS (learning management system) that will be used to access the pre-class content, which includes a video lecture, digital reading passages, and interactive tasks. This hands-on training session instructed the student on how to use the LMS, download resources, and perform pre-class activities. In addition, students engaged in and practiced a sample of the pre-class content, as well as a mock task in class, to learn how applying that knowledge would be assisted in real-world collaborative efforts. Teacher support was offered during the training session, and participants were also permitted to ask questions and express any concerns they had in the days leading up to the training week to help them feel more at ease with the flipped approach. The primary purpose of the training step was to ensure that the students were comfortable with and capable of using the technology for the flipped classroom intervention.

Week 2: Pre-Intervention Assessment

A full language skills pretest was administered to both the experimental and control groups. Furthermore, the demographic information and baseline data on technology use and orientation sessions were collected, and the study protocols and what students could expect were also explained.

Weeks 3-6: Instructional Intervention (Four Weeks)

(a). Experimental Group (Flipped Classroom Model)

The experimental group engaged in instruction using the flipped classroom model across four instructional units. Each unit focused on the integrated development of the four language skills:

Unit 1: Reading Skill Focus

In this unit, pre-class activities included viewing video lectures explaining reading strategies and vocabulary and engaging with digital reading passages with comprehension questions, while in-class activities included small group discussions analyzing texts, corresponding comprehension tasks, applying strategies with teacher assistance (for example, modeling practices), and peer teaching, among other activities. Additionally, students also received individualized feedback on readings.

Unit 2: Writing Skill Focus

Pre-class activities included students reviewing video tutorials on how a paragraph is organized and writing conventions. They also looked at model essays and writing samples, and they engaged in pre-writing organizational tasks.

Meanwhile, in-class activities included teacher-directed collaborative brainstorming, a peer review workshop, teacher-guided writing with instant feedback, revising activities, and individual conferences to address writing issues.

Unit 3: Listening Skill Focus

Unit 3 included a pre-class activity where students listened to audio recordings with a guide, a listening task via a video lecture on listening strategies, and a prepared vocabulary list for listening. The in-class activity focused on listening practice with an authentic task strategy, applying task discussion about the listening challenge, collaborative note-taking practice, and target skill development tasks.

Unit 4: Speaking Skill Focus

Pre-class activities included viewing video examples of effective oral communication, listening to sample conversations and presentations, and preparing a strategy for speaking by making outlines, while in-class activities included teacher-directed speaking practice activities, role-playing activities, group discussions and debates, presentation tasks with feedback from peers and the instructor, pronunciation clinics, and fluency practice activities.

During the intervention period, students in the experimental group completed assignments via a learning management system. Class sessions were designed to maximize time for interaction and student learning, with the teacher facilitating, giving feedback, and helping students address difficulties while working with each other.

(b). Control Group (Traditional Instruction)

Participants in the control group received instruction using traditional methodologies. These included teachers conveying instructional content during class time through teacher demonstrations and explanations. Students also listened to the teacher's presentations, took notes, and participated in teacher-led activities. Moreover, practice activities were assigned as independent homework; class time included very little interactive practice with the target forms; and feedback was primarily through homework grading and periodic tests.

The main distinction between the control and experimental groups was in the instructional approach and the organization of activities, while learning objectives and curricular content, which addressed the four language skills, were all equivalent in both groups.

Week 7: Post-Intervention Assessment

In Week 7, teachers administered the post-test of language skills to the two groups, administered the questionnaire on student engagement to both groups, and administered the satisfaction assessment to participants in the experimental group. Teacher observations and notes on implementation were also collected.

C. Data Collection and Analysis

Data gathering took place in the second term of the 2024-2025 academic year at Fatima Al Zahraa Secondary School in Alexandria, Egypt. The data collection phase included the three main instruments used in both the experimental and control groups in this study. Additionally, a pretest and posttest measured the overall proficiency in the English language by assessing the four language skills of reading, writing, listening, and speaking. A questionnaire to assess student engagement in the behavioral, emotional, and cognitive aspects was also provided after the intervention, and a satisfaction survey was provided only for the experimental group to get their perceptions on their learning experience.

These instruments were administered in controlled environments in the same manner for both groups. Students completed the measurement tools independently (i.e., without help), and participation was voluntary, as an appropriate informed consent protocol was observed. Students were not graded based on the research assessments to reduce anxiety towards performance pressure, encouraging honest responses.

TABLE 1
DATA COLLECTION TIMELINE

Week	Activity	Instrument Used	Groups
0	Training Stage for the Flipped Classroom Model	Orientation and LMS training	Experimental Only
1	Pre-Test Administration	Comprehensive Language Skills Test (100 points total)	Both
2-5	Instructional Intervention	Flipped vs. Traditional Instruction	Both
6	Post-Test Administration	Comprehensive Language Skills Test (same structure)	Both
6	Engagement Assessment	Student Engagement Questionnaire (25 items)	Both
6	Satisfaction Survey	Student Satisfaction Survey (10 items)	Experimental Only

The research team employed the Statistical Package for the Social Sciences (SPSS) version 28 for thorough data analysis. The variables were coded as follows:

The pretest total score was based on a scale of 0 to 100 points.

- Reading subscale (0-25)
- Writing subscale (0-25)
- Listening subscale (0-25)
- Speaking subscale (0-25)

Post-test total scores also ranged from 0 to 100 points with the same subscale structure. Engagement Scores (Scale: 25-125 Points):

- Behavioral Engagement Subscale

- Emotional Engagement Subscale
- Cognitive Engagement Subscale

After completing data entry and verifying its completeness and accuracy, statistical analyses continued using multiple analytic approaches:

(a). *Paired Sample T-Test*

A paired sample t-test was performed to estimate the differences in pretest and posttest scores between the control and experiment groups to identify any significant learning improvements for both groups after the intervention period.

(b). *Independent Sample T-Test*

An independent sample t-test was performed to examine posttest scores and compare performance between the two groups to see if flipped classroom learning produced substantially different results than traditional learning.

(c). *Analysis of Covariance (ANCOVA)*

An Analysis of Covariance (ANCOVA) was used to provide statistical analyses, with pretest scores serving as covariates to compensate for existing variations and provide a more accurate estimate of the intervention's impact on posttest performance.

(d). *Descriptive Statistics*

Descriptive statistics, such as averages, standard deviations, and percentages, were used to summarize patterns in overall performance, levels of involvement, and responses to closed-ended questionnaire items.

(e). *Effect Size Calculation*

Cohen's d effect sizes were calculated to highlight the practical significance of observed differences and to shed light on the effect size in terms of the practical relevance of intervention effects when differences are accounted for beyond statistical significance. Statistical significance was evaluated for all inferential analyses at a level of 0.05. When determining and interpreting outcomes from the deployment of the flipped classroom intervention, practical importance, as shown by the effect size, was considered alongside statistical significance.

(f). *One-Way ANOVA Analysis*

A one-way analysis of variance (ANOVA) was used to investigate the relationship between language skill development and engagement, specifically whether the engagement score (behavioral, emotional, and cognitive) reliably predicted the posttest language proficiency score (reading, writing, listening, and speaking level) in the experimental flipped classroom and control group. The engagement score was categorized as high, medium, and low based on the tertile split of the overall engagement score (25-125). The Tukey HSD post hoc analysis was used to find the source of the differences between groups. The analysis also considers pretest scores to control differences in the baseline language proficiency.

IV. RESULTS AND DISCUSSION

A. *Results of Overall Language Proficiency Achievement*

To evaluate the total effect of flipped classroom instruction on overall student English language proficiency, comprehensive pre- and posttest English language proficiency achievement scores yielded from both the experimental and control groups were analyzed. Before and after the instructional intervention, within-group paired sample t-tests compared the total language proficiency scores, and the outcomes of the t-test analyses are presented in Table 2.

TABLE 2
PAIRED SAMPLES T-TEST OVERALL LANGUAGE PROFICIENCY RESULTS

Group	Test	Mean Score (out of 100)	Standard Deviation	t-value	p-value
Experimental	Pre-Test	56.23	8.45		
	Post-Test	78.67	7.12	14.832	0.000*
Control	Pre-Test	55.87	8.72		
	Post-Test	62.43	8.29	4.561	0.000*

Note: *indicates significance at $p < 0.01$ level.

The pre-test results show that both the control and the experimental group had similar levels of English language proficiency before the instructional intervention. The experimental group had a mean score of 56.23 (SD = 8.45), and the control group had a mean score of 55.87 (SD = 8.72). An independent samples t-test confirmed that there was no significant difference between the groups at pretest ($t = 0.168$, $p = 0.867$), confirming that at pretest the two groups were equal and enabling meaningful comparison of the groups after the intervention.

After the six-week instructional intervention was over, both groups showed statistically significant growth in overall language proficiency levels. The experimental group, who received instruction in a flipped classroom model, experienced greater gains with a mean overall language proficiency post-test score of 78.67, SD 7.12, which was a 22.44-point increase from the pre-test and a 39.9 percent increase from the pre-test. The paired sample t-test demonstrated a statistically

significant difference ($t=14.832$, $p=0.001$), suggesting that the observed change was not due to insignificant variabilities. The effect size of the intervention (Cohen's $d=2.87$) indicated that the finding had a large practical importance.

The control groups, who were taught using traditional instruction only, also demonstrated improvement that was statistically significant, as indicated by a posttest mean of 62.43 and an SD of 8.29, which translated to an increase of 6.56 points, an 11.7 percent increase from the baseline test ($t=4.561$, $p=0.001$). This improvement serves as evidence that students were, indeed, learning as indicated by the standard instruction. However, the increases in the control group were of much lower magnitude than the increases of the flipped instruction group, with a moderate effect size (Cohen's $d=0.79$).

These findings are in line with previous research literature surrounding flipped classroom approaches to language learning. For example, Öztürk and Çakıroğlu (2021) observed a trend of increased language development and attributed it to opportunities for learner-led active learning, the increased autonomy of learners in their learning, and greater opportunities for enacting quality practice (Hung, 2015).

An independent samples t-test was conducted to technically examine the effectiveness of flip instruction against the standard instruction group based on post-test groups. Post-test comparability results are illustrated in Table 3.

TABLE 3
INDEPENDENT SAMPLES T-TEST OF POST-TEST COMPARABILITY BETWEEN GROUPS

Group	Mean Post-Test Score	Standard Deviation	t-value	p-value	Cohen's d
Experimental	78.67	7.12	8.234	0.000*	2.13
Control	62.43	8.29			

Note: *specified significance at the $p < 0.01$ level.

The descriptive statistics showed that the experimental group obtained significantly higher posttest mean scores ($M=78.67$) compared to the control group ($M=62.43$), with the magnitude of mean difference being 16.24 points. The independent t-test turned out also very significant ($t=8.234$, $p<0.001$), which can be considered as a statistically significant difference, and the researchers are sure that the difference is due to the pretest/posttest procedure and is not just random. The practical size of the effect (Cohen $d=2.13$) indicated that the flipped classroom intervention was not only statistically but also practically significant for both groups, as far as the findings in their language learning and acquisition implied. The study's quantitative results revealed that employing the flipped classroom improved EFL learners' English language skills (Najmi, 2020). Numerous pedagogical factors can explain what the experimental group achieved in the study from using the pedagogic frame of the flipped classroom.

First, the application of the flipped classroom strategy provided more time for active learning in class as content delivery took place outside of class, and consequently, the students devoted more time to practicing a wide range of language skills. Second, the flipped strategy enables students to absorb the instructional input out of class at their own pace and to access multiple views of the content if necessary to reach full understanding. Each person has a unique approach to dealing with and counteracting cognitive overload while learning the skill. Third, since students are already engaged in active learning, flipped instruction increases the effectiveness of the third phase—immediate evaluation and face-to-face remedial instruction. Finally, the flipped method encourages collaborative learning and interaction with peers in the target language, thus creating a more communicative and social learning environment, abundant with collaborative learning and meaningful interaction (Evseeva & Solozhenko, 2015). Implications of this study are consistent with theoretical approaches that emphasize the significance of input, output, and interaction for second language acquisition.

B. Results Related to Individual Language Skills

To describe more fully the individual intervention effects of the flipped classroom on students' language skills, separate analyses were conducted for each of the four language skills, which are represented in the comparative results between the experimental and control groups in Table 4: reading, writing, listening, and speaking skills.

TABLE 4
POST-TEST PERFORMANCE BY INDIVIDUAL LANGUAGE SKILL

Language Skill	Group	Mean Score (out of 25)	Standard Deviation	t-value	p-value	Cohen's d
Reading	Experimental	19.87	2.14	6.452	0.000*	1.67
	Control	15.73	2.48			
Writing	Experimental	19.23	2.35	7.128	0.000*	1.84
	Control	14.90	2.31			
Listening	Experimental	20.13	1.89	5.891	0.000*	1.52
	Control	16.47	2.43			
Speaking	Experimental	19.44	2.28	6.734	0.000*	1.74
	Control	15.33	2.37			

Note: *shows significance at the $p < 0.01$ level.

As presented in Table 4, the students in the experimental group achieved higher scores than the students in the control group across all language skills.

(a). Reading Comprehension

Students in the experimental groups achieved better scores than the control group in reading comprehension ($M = 19.87$, $SD = 2.14$ vs. $M = 15.73$, $SD = 2.48$), $t(65) = 6.452$, $p < 0.001$, $d = 1.67$. The inverted classroom provided students with the chance to practice reading by letting them read assigned reading texts and practice reading skills outside of class before engaging in deeper excitement and comprehension objectives in class. This was consistent with the academic literature demonstrating the need for additional engagement with the texts and also reading strategy practice (Lee & Wallace, 2018).

(b). Writing Proficiency

The experimental group had significantly better writing abilities ($M=19.23$, $SD=2.35$) than the control group ($M=14.90$, $SD=2.31$) and reported significant differences at 7.128 and $p=0.001$. Moreover, the significant effect size was 1.84. Many of these were aligned with Fathi and Rahimi (2022), who noted significant gains in complex writing accuracy and fluency with flipped instruction. The students had ample preparation before class and then had enough time to write in class through peer review, and the teacher's feedback presented ideal conditions for writing development (Al-Harbi & Alshumaimeri, 2016).

(c). Listening Comprehension

The student in the experimental group had a higher listening score ($M=20.13$, $SD=1.89$) than the students in the control group ($M=16.47$, $SD=2.43$) for a significant effect at 5.891 and $p=0.001$, with a large effect size of 1.52. This finding appears to support the earlier work of Qiu and Luo's (2022) report, which enhances listening performance and decreases listening anxiety when doing flipped classroom practice. The potential to preview listening materials, develop specific listening strategies, and receive support during in-class activities resulted in improved listening performance (Gok et al., 2021).

(d). Speaking Skills

The experimental group showed significantly higher speaking, achieving $M=19.44$ and $SD=2.28$, whereas the control group attained $M=15.33$ and $SD=2.37$, demonstrating a significant difference at 6.734 and $p=.001$ and a large effect size of 1.74. These results were consistent with Teng (2018), who found improvement in oral proficiency after flipping instruction. The flipped instruction models promote speaking development by freeing up class time for speaking practice while lowering cognitive load by not processing new content and producing output in the same time period (Najmi, 2020).

The replicable pattern of superiority in performance across all four language skills provides further supporting evidence that the flipped classroom model meets the holistic language development needs of EFL learners. The flexibility of the approach to support the simultaneous development of multiple skill areas is a notable advantage of the pedagogical model, given that integrated skills' development reinforces communicative competence (Mehring, 2016).

C. Student Engagement Results

To meet the goal of the research in relation to student engagement, participants were administered a comprehensive post-intervention questionnaire, which measured behavioral, emotional, and cognitive engagement for both groups. Table 5 displays overall engagement results and subscale comparisons.

TABLE 5
STUDENT ENGAGEMENT SCORES BY GROUP AND DIMENSION

Engagement Dimension	Group	Mean Score	Standard Deviation	Maximum Possible	%	t-value	p-value
Overall Engagement	Experimental	102.47	8.23	125	82.0%	9.156	0.000**
(25 items)	Control	78.93	9.45	125	63.1%		
Behavioral Engagement	Experimental	37.23	2.87	45	82.7%	8.734	0.000**
(9 items)	Control	28.67	3.12	45	63.7%		
Emotional Engagement	Experimental	33.17	2.94	40	82.9%	9.012	0.000**
(8 items)	Control	25.43	3.28	40	63.6%		
Cognitive Engagement	Experimental	32.07	3.15	40	80.2%	8.567	0.000**
(8 items)	Control	24.83	3.41	40	62.1%		

Note: **indicate significance at the $p<0.01$ level.

The engagement analysis showed that the student in the experimental condition was significantly more engaged in overall engagement than the student from the control group ($M=102.47$, 82.0%, and $M=78.93$, 63.1% points), $t=9.156$, $p<0.001$, a difference of 18.9 percentage points. This major difference in level of engagement suggests that the flipped classroom model supports a more engaging learning experience that engages students with more cognitive affection and behavior in learning English.

(a). Behavioral Engagement

The experimental group showed significantly higher levels of behavioral engagement ($M=37.23$, 82.7%) compared to the control one ($M=28.67$, 63.7%), $t=8.734$, $p<0.001$. It was suggested that students in flipped classes were more engaged behaviorally, showed a higher attendance rate, completed assignments more often, and participated in class interaction more frequently while showing a higher level of on-task behavior. The teacher also notes that students in the experimental

condition come to class more prepared, participate with more enthusiasm during the activity, and put in strong effort into completing the learning task (Ardi & Rianita, 2022).

(b). *Emotional Engagement*

Students enrolled in a flipped instructional model indicated significantly greater levels of emotional engagement ($M = 33.17$, 82.9%) than participants following a traditional instructional model ($M = 25.43$, 63.6%). $t = 9.012$, $p < 0.001$, indicating the flipped instructional model engaged students more positively in the learning of English, in terms of positive feelings towards learning English, interest in course content, enjoyment of learning activities, and identification with being an English learner. Qualitative feedback indicated students experienced the flipped instruction as more pleasurable, less stressful, and more personally valuable than the traditional instructional model (Huseinović, 2024).

(c). *Cognitive Engagement*

The experimental group indicated significantly greater cognitive engagement ($M = 32.07$, 80.2) than the control group ($M = 24.83$, 62.1). $t = 8.567$, $p < 0.001$, demonstrating that students in the flipped classroom approached learning from a deeper learning style, demonstrated metacognitive awareness, displayed greater mental effort in the learning task, and approached challenging learning materials with much perseverance. The findings for engagement corroborate findings from Öztürk and Çakıroğlu (2021).

In the link between engagement and language skill development, the one-way ANOVA revealed a significant effect for engagement on post-test language proficiency, $F(2,57)=12.45$, $p<0.001$, $\eta^2=0.31$. The post hoc Tukey HSD test showed that there was much higher post-test proficiency for the high-engagement experimental group ($M=82.35$, $SD=6.78$) when compared with both the medium-engagement ($M=75.12$, $SD=7.23$) and low-engagement ($M=68.45$, $SD=8.10$) experimental groups in the post-test for each language skill ($p<0.05$). There were no significant differences between the engagement groups in the control group ($p>0.05$). This finding suggested that engagement—especially in the flipped classroom model—is an important predictor for growth in language skill, and it supported the hypothesis that engagement mediates the effectiveness of flipped instruction.

TABLE 6
HIGH-RANK STATEMENTS FROM EXPERIMENTAL GROUP

Statement	Mean Score (Out of 5)	Percentage Agreement (4-5 on Scale)
The flipped classroom approach helped me learn English more effectively.	4.57	93.3%
I enjoyed having access to learning materials that I could review at my own pace.	4.53	90.0%
Class time activities were more interesting and engaging than in traditional classes.	4.47	90.0%
I felt more confident participating in class activities after preparing at home.	4.43	86.7%
The flipped model helped me develop better study habits and learning strategies.	4.37	86.7%

Responses to the survey offer a strong indication that the flipped classroom has been viewed very positively by participants. In the experimental group, the most highly rated statement, “The flipped classroom helps me learn English better” ($M=4.57$ or 93.3% agreement), indicated that students view the instructional method as successful.

Students greatly appreciated learning in a flipped classroom, especially because of the demonstrated flexibility and agency it offered. For example, there was strong agreement with “I enjoy having access to learning materials, so I reviewed them at my own pace” ($M=4.53$, or 90.0% agreement). As mentioned, this flexibility will enable learning variety that considers individual student variances in engagement, access, and interaction with academic content (Azzouz Boudadi & Gutiérrez-Colón, 2020).

The statement regarding class activities being more engaging ($M = 4.47$, or 90.0% agreement) also suggested that this reformulated use of class time creates a more engaging and captivating method for learning, while the use of class time for in-class interactive practice and participation in learning activities, rather than receiving information passively, was making a difference in classroom environments that participants found engaging, motivating, and fun (Ardi & Rianita, 2022). More importantly, students had indicated increased confidence in classroom participation after home preparation ($M=4.43$, or 86.7% agreement), indicating that the flipped model contributed to reducing anxiety during performance and arriving better prepared.

The students also noted that the flipped model had facilitated the development of meaningful learning strategies and study skills ($M=4.37$, and 86.7% agreement). This metacognitive awareness and strategy development was another added benefit beyond the relatively immediate language skill benefits and contributed to eventual long-term learning success. This enhancement of engagement outcomes, combined with better levels of achievement, will add additional evidence in supporting the flipped learning of EFL learners (Huseinović, 2024).

V. CONCLUSION

This is a strong empirical study that offers convincing proof for endorsing the flipped instruction model in short-term English proficiency and involvement by secondary EFL students. According to the study results, the flipped instruction method had specific benefits compared to the traditional instruction in many learning outcomes. Also, the result of engagement showed much improvement from the use of the flipped approach. This positive impact is attributed to the

following points:

The first one is the possibility for the flipped model to optimize learning potential by active learning chance, removing content delivery during instruction time, and providing the instruction time to practice meaningful communication and apply skills in face-to-face interaction. This sequence of learning time also aligns with the research results that show the process of language learning is a developing process requiring time and engagement with active language usage rather than just passive language use (Evseeva & Solozhenko, 2015). Second, there exists in the flipped model the opportunity for managing personal differences in learning, depending on how the student is capable of going through his/her instructional material at his/her own pace. This involves the capability of going through content as many times as one can to achieve the realization. Students can learn at their own rhythm and at a much slower pace than the teacher and can completely and thoroughly process the input of the lingo without being distracted by anxiety about not being able to follow the teacher (Mehring, 2016).

Third, flipped teaching allows one to offer timely feedback and remedial learning at an essential phase of learning the skill application. In particular, within the classroom setting, the teachers can observe the performance and identify any challenges as they arise and deliver corrective instruction at the very point that students are working on language skills application (Başal, 2015).

Fourth, the inverted classroom provides an opportunity for collaborative learning, peer-to-peer learning, and communicative use of the language, which is frequently extremely difficult to imitate in the traditional environment. The pedagogical structure of content delivery duration would enable the teacher with the power to structure active group work, pair work, conversations, and other interactive activities, all of which would be valuable interactive experiences capable of supporting rich interaction and social learning. Such collaborative interactions are also directly concerned with the social nature of transformational language learning and provide real-life contexts for creating communicative competence (Azzouz Boudadi & Gutiérrez-Colón, 2020).

Finally, fifth point implies that it promotes autonomy, self-regulated learning among students, and metacognitive awareness. Students gain the ability to create their own learning, keep track of their knowledge, evaluate their progress, and correct their approaches to learning: an ability that they will not only find useful in their current learning success but also allow them to obtain and develop the skills of such activity in the future (Nikmah, 2020).

A. Limitations of the Study

While the researchers provide useful information to support the efficacy of flipped classes, numerous limitations must be addressed. The study was conducted in a single geographical location and included a limited sample of students from a single secondary school, limiting the study's findings. The intervention lasted six weeks, which was long enough to see substantial impacts from the intervention in the trial; however, this length did not adequately capture the durability of improvement after the study or any delayed effects on learning.

B. Recommendations and Pedagogical Considerations

The researcher suggests some recommendations for practice in education in view of the ample evidence highlighting the flipped classroom's effectiveness provided in this study. First, flipped classrooms should be considered as a guiding model for EFL programs in schools, particularly in secondary and post-secondary educational contexts, where students develop an essential level of technological literacy and self-regulated learning (Mehring, 2016). Second, professional learning development for teachers should include content on flipped instruction principles and how to work with technology and create effective content materials and design to implement that practice (Akçayır & Akçayır, 2017). Third, schools must develop the appropriate technological infrastructure to support flipped instruction, including internet access, learning management systems, and content creation tools. Fourth, curriculum developers should create well-designed, quality pre-class materials, such as instructional videos, interactive exercises, and digital resources for actively engaging learners, to represent content clearly and provide multiple access points and the ability to access content quickly (Başal, 2015). Fifth, teachers should create engaging classroom activities that maximize the time for active learning, promote collaboration and interaction, and provide opportunities for authentic language practice for all four skills (Ardi & Rianita, 2022). Finally, the implementation should be gradual and mapped out, carefully starting with a pilot program, obtaining student feedback, making refinements based on experiences, and incrementally expanding with systematic models (Evseeva & Solozhenko, 2015).

REFERENCES

- [1] Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1-11. <https://doi.org/10.1016/j.edurev.2016.11.002>
- [2] Al-Harbi, S., & Alshumaimeri, Y. (2016). The flipped classroom impact in grammar class on EFL Saudi secondary school students' performances and attitudes. *English Language Teaching*, 9(10), 60. <https://doi.org/10.5539/elt.v9n10p60>
- [3] Alsalam, A. I. (2025). Flipped learning in EFL classrooms: A systematic review of implementation and outcomes. *Journal of King Abdulaziz University-Arts and Humanities*, 33(5), 581 - 607. <https://doi.org/10.4197/Art.33-5.20>
- [4] Ardi, P., & Rianita, E. (2022). Leveraging gamification into EFL grammar class to boost student engagement. *Teaching English with Technology*, 22(2), 90-114.
- [5] Azzouz Boudadi, N., & Gutiérrez-Colón, M. (2020). Effect of gamification on students' motivation and learning achievement in

- second language acquisition within higher education: A literature review 2011-2019. *The EUROCALL Review*, 28(1), 57-69. <https://doi.org/10.4995/eurocall.2020.12974>
- [6] Başal, A. (2015). The implementation of a flipped classroom in foreign language teaching. *Turkish Online Journal of Distance Education*, 16(4). <https://doi.org/10.17718/tojde.72185> (researchgate.net)
 - [7] Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.
 - [8] Chen Hsieh, J. S., Wu, W.-C. V., & Marek, M. W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, 30(1-2), 1-21. <https://doi.org/10.1080/09588221.2015.1111910>
 - [9] Eltahir, M. E., & Alsalthi, N. R. (2025). Impact of the flipped classroom on academic achievement, motivation, and engagement: A higher education case study. *Contemporary Educational Technology*, 17(1), ep553.
 - [10] Evseeva, A., & Solozhenko, A. (2015). Use of flipped classroom technology in language learning. *Procedia – Social and Behavioral Sciences*, 206, 205-209. <https://doi.org/10.1016/j.sbspro.2015.10.006> (researchgate.net)
 - [11] Fathi, J., & Rahimi, M. (2022). Examining the impact of flipped classroom on writing complexity, accuracy, and fluency: A case of EFL students. *Computer Assisted Language Learning*, 35(5-6), 1188-1206. <https://doi.org/10.1080/09588221.2020.1825097>
 - [12] Gok, D., Bozoglan, H., & Bozoglan, B. (2021). Effects of online flipped classroom on foreign language classroom anxiety and reading anxiety. *Computer Assisted Language Learning*, 1-21. <https://www.tandfonline.com/doi/full/10.1080/09588221.2021.1950191>
 - [13] Hung, H.-T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning*, 28(1), 81–96. <https://doi.org/10.1080/09588221.2014.967701> (dacemirror.sci-hub.box)
 - [14] Huseinović, L. (2024). The effects of gamification on student motivation and achievement in learning English as a foreign language in higher education. *MAP Education and Humanities*, 4, 10-36.
 - [15] Khaleel, F. L., Ashaari, N. S., & Wook, T. S. M. T. (2020). The impact of gamification on students' learning engagement. *International Journal of Electrical and Computer Engineering (IJECE)*, 10(5), 4965-4972. <https://doi.org/10.11591/ijece.v10i5.pp4965-4972>
 - [16] Lee, G., & Wallace, A. (2018). Flipped learning in the English as a Foreign Language classroom: Outcomes and perceptions. *TESOL Quarterly*, 52(1). <https://doi.org/10.1002/tesq.372> (researchgate.net)
 - [17] Lo, C. K., & Hew, K. F. (2017). A critical review of flipped classroom challenges in K-12 education: Possible solutions and recommendations for future research. *Research and Practice in Technology Enhanced Learning*, 12(1), 4. <https://doi.org/10.1186/s41039-016-0044-2>
 - [18] Mehring, J. (2016). Present research on the flipped classroom and potential tools for the EFL classroom. *Computers in the Schools*, 33(1), 1-10. <https://doi.org/10.1080/07380569.2016.1139912>
 - [19] Mizza, D., Reese, M., & Malouche, D. (2025). Flipped classroom evaluation and blended learning potential: A case study of engagement and inclusion in quantitative education. *Smart Learning Environments*, 12, Article 56.
 - [20] Najmi, A. H. (2020). The effectiveness of flipped classroom approach on students' achievement in English language in Saudi Arabian Southern Border Schools. *International Education Studies*, 13(9), 66. <https://doi.org/10.5539/ies.v13n9p66> (ccsenet.org)
 - [21] Nikmah, H. (2020). Gamification to improve students' engagement in learning English. *ACITYA Journal of Teaching & Education*, 2(1), 60-70. <https://doi.org/10.30650/ajte.v2i1.277>
 - [22] Oraif, I., & Elyas, T. (2021). The impact of COVID-19 on learning: Investigating EFL learners' engagement in online courses in Saudi Arabia. *Education Sciences*, 11(3), 99. <https://doi.org/10.3390/educsci11030099>
 - [23] Öztürk, M., & Çakıroğlu, Ü. (2021). Flipped learning design in EFL classrooms: Implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments*, 8, Article 2. <https://doi.org/10.1186/s40561-021-00146-x>
 - [24] Qiu, Y., & Luo, W. (2022). Investigation of the effect of flipped listening instruction on the listening performance and listening anxiety of Chinese EFL students. *Frontiers in Psychology*, 13, Article 1043004. <https://doi.org/10.3389/fpsyg.2022.1043004>
 - [25] Teng, M. F. (2018). Flip your classroom to improve EFL students' speaking skills. In *Innovations in flipping the language classroom* (pp. 113-122). Springer. https://doi.org/10.1007/978-981-10-6968-0_9

Rasha A. Dorgham, Associate professor of Applied Linguistics/TESOL, College of Languages and Translation, Imam Mohammed Bin Saud Islamic University (IMSIU) / Lecturer, Cairo University, Faculty of Graduate Studies for Education. Her areas of expertise include second language acquisition and assessment, computer-assisted language learning, teaching effectiveness, and teacher education. Her research interests lie in the areas of second language teaching and learning, language assessment, and teacher cognition.

Lamya Bin Obiad, Assistant professor of Applied Linguistics/TESOL, College of Languages and Translation, Imam Mohammed Bin Saud Islamic University (IMSIU) / Lecturer. Her areas of expertise include second language acquisition and assessment, computer assisted language learning, teaching effectiveness, and teacher education. Her research interests lie in the areas of second language teaching and learning, language assessment, and teacher cognition.

Media Discourses of Moderation: A Critical Discourse Analysis of Global Coverage of Ahmed Al-Sharaa's Presidency

Ibrahim Bani Abdo

Department of English Language and Literature, University of Jordan, Jordan

Abdelhameed Al Awabdeh

Department of English Language and Literature, University of Jordan, Jordan

Mohammad Albtoush

Department of English Language and Literature, University of Jordan, Jordan

Abeer Hameed Albashtaw

Department of English, Luminus Technical University College, Jordan

Nabil Al-Awawdeh

Department of Modern Languages, Yarmouk University, Jordan

Mheel Al-Smaihyeen

Department of English Language and Literature, University of Jordan, Jordan

Abstract—This article examines how global media portrayed the Syrian President Al-Sharaa before and after he assumed the presidency. It employs critical discourse analysis to investigate how media coverage influences the public perceptions of him and the nature of political transition during these two periods. The study uses Fairclough's model; the Van Leeuwen framework; and evaluative theory to analyze the linguistic techniques used in the global media. Through analysis of fifty English-language articles published in leading international newspapers between 2020 and 2025, this study identifies four key conclusions: First, the news media portrays a significant shift of Al-Sharaa's intellectual and political trajectory. Second, it employs laudatory language that presents him as 'open', 'pragmatic', and 'prone to change', which fosters a more appealing image for the audience. Third, it depicts Al-Sharaa's language as actively shaping events rather than passively responding to events. Finally, it links Al-Sharaa's discourse to concepts of national cohesion and community service embodying a political transformation with collective implications, even though this transcendence is influenced more by personal factors. The results reveal that the media not only serves as a reporting entity but actively participates in reshaping the political landscape by the manner in which it portrays and reproduces leaders' images. The study analyzes only English-language content. Future investigations should include multiple languages, examine public reactions, and compare Al-Sharaa's image with other leaders undergoing similar shifts. In conclusion, the research verifies that media discourse acts as a symbolic force shaping leader portrayals and public opinion.

Index Terms—global media, public perceptions, linguistic techniques, political transformation

I. INTRODUCTION

In contemporary times, the media has become the primary gateway through which public perceptions of political leaders are reshaped. Published materials, whether in the form of reports, analyses, or digital content, do not merely convey news; they actively participate in constructing political meaning and directing public interpretation of events. Fairclough (1995) highlights this idea, noting that media discourse selects certain meanings to emphasize while neglecting others, leading to a restructuring of how audiences understand political reality. Van Dijk (2008) argues that media messages possess the power to reconstruct collective perceptions and reshape political affiliations and identities. Therefore, studying the mechanisms by which leaders' images are shaped within media discourse is crucial, especially during periods of political upheaval or widespread debate.

Within this framework, Ahmed Al-Sharaa's recent presence in international media has emerged as an example of a leader being reinterpreted in a different light. Many news outlets have portrayed him as a figure who has abandoned a harsh rhetoric and adopted more balanced and open positions. This discourse of "transition" has become a constant element in the media narrative surrounding him. This pattern of approach opens the door to several questions: What linguistic and narrative mechanisms do the media employ to construct a story of political transition from a hardline tone

to a more moderate approach? And what impact does this narrative have on shaping how others perceive him? Such characterization may alter the way international actors engage with the leader and reshape the interpretation of his policies and decisions. Analyzing this discourse also allows for a deeper understanding of how concepts like moderation, reform, and extremism are circulated in the global media, and how these terms are transformed into tools for shaping the political narrative in the current era.

Although Al-Sharaa's political transformation has received extensive international media coverage, academic output examining how such transformations are presented or analyzed from a linguistic perspective remains extremely limited. The vast majority of studies in political discourse focus on leaders from the Western context, while the representation of leaders in the Arab world and the Middle East receives far less research attention. Consequently, our understanding of how the global media frames the concept of "moderation" when discussing Al-Sharaa remains incomplete.

A review of the literature also reveals a clear absence of studies employing critical discourse analysis to deconstruct how global media portray Al-Sharaa and his transition from moving from a hardline stance to a more moderate approach. This deficiency is further highlighted by the scarcity of research comparing the framing techniques used across various international media platforms. In light of this research gap, this study aims to fill the gap by analyzing the discursive representations and framing mechanisms employed in global media coverage of Al-Sharaa. Hence, the study begins to answer the following questions:

- i. How do international media outlets use their linguistic resources to portray Al-Sharaa's shift from hardline rhetoric to a more moderate stance?
- ii. What rhetorical strategies do global media outlets employ to present Al-Sharaa's presidency as a form of political transformation?

II. LITERATURE REVIEW

A. *Media Discourse and the Construction of Political Meaning*

Media plays a pivotal role in shaping how people perceive political figures and public issues. Several researchers argue that the media's function extends beyond simply reporting facts; it encompasses the construction and reconstruction of meaning. According to Fairclough (1995), media outlets selectively highlight certain events while neglecting others, thereby reshaping public perception of the political landscape. Van Dijk (2008) demonstrates that the media relies on intellectual and ideological frameworks that influence how the audience interprets political actions. Hall (1997) adds that media representation is not neutral and that editorial choices are a key factor in shaping public opinion.

By using specific headlines, highlighting recurring themes, employing particular images, and relying on words with specific connotations, the media can shape a mental image of leaders, portraying them as moderates, potential threats, reformers, or authoritarian rulers (Machin & Mayr, 2012). This influence is amplified during periods of significant political change or crises.

B. *Extremism, Moderation, and the Construction of Political Impressions*

The literature suggests that concepts like "extremism" and "moderation" are not neutral labels, but rather are produced through discourse and reinforced by repeated use in media and politics. Schmid (2013) argues that these terms are socially constructed through media platforms that recycle and define their meanings. Busher (2019) further explains that political actors employ these classifications to serve their interests, either by presenting themselves positively or by stigmatizing and delegitimizing other parties. Tajfel's (1981) social classification theory explains how these linguistic categories contribute to the formation of collective identities that shape public opinion.

When a leader is labeled "moderate," the public tends to interpret their actions more favorably and acceptably, even if their political practices do not change significantly (Lakoff, 2002; Chong & Druckman, 2007). This pattern is particularly evident in the Middle East, where some international media outlets resort to simplistic narratives to describe complex political transformations (Lynch, 2016; Pintak, 2014). This approach to portraying leaders can have a tangible impact on the course of diplomacy, global public opinion, and the political legitimacy of leaders (Said, 1997; Hammond, 2017).

C. *Critical Discourse Analysis (CDA) and Political Leadership*

Critical discourse analysis is one of the most important tools used to understand how the image of political leaders is constructed within the media. This approach focuses on uncovering the hidden relationships between language, power, and ideology, and the meanings implicitly conveyed through media texts (Fairclough, 2010; Wodak, 2011). Wodak and Meyer (2016) indicate that this type of analysis enables researchers to trace the deep structures that shape media discourse, including framing techniques that determine how the leader is presented to the public (Entman, 1993), lexical choices used to convey connotations of danger, trust, influence, or reform (Richter, 2021a, 2021b), and narrative structures employed to construct narratives of change, renewal, or conflict (Van Leeuwen, 2008). Critical discourse analysis also reveals the implicit dichotomies upon which the media relies, such as moderation versus extremism or stability versus chaos (Chouliaraki & Fairclough, 1999). Thanks to this combination of analytical tools, critical discourse analysis becomes a suitable framework for understanding how global media is reshaping the image of Al-Sharaa as a leader moving from a hardline rhetoric to a more moderate stance.

D. Global Media Representation of Middle Eastern Leaders

A wide range of research indicates that international media often adopt repetitive narrative patterns when reporting on political leaders in the Middle East. Said (1997) argues that Western media sometimes relies on preconceived cultural images that result in a simplistic or biased reading of political reality. Hammond (2012) adds that global media tends to categorize regional leaders into pre-defined stereotypes such as “reformers,” “hardliners,” or “authoritarian rulers.” Recent studies also reveal that international coverage relies on recurring, stereotypical themes—such as security, stability, conflict, or reform—as interpretive keys to understanding events in the Middle East (Pintak, 2014; Lynch, 2016). These media practices highlight the need to analyze how the media constructs the concept of “moderation” when discussing political figures like Al-Sharaa.

E. Recent Studies on Political Transformation in Media Discourse (Elaborated)

Although studies focusing directly on Al-Sharaa are scarce, recent literature illuminates how the media constructs narratives of political transformation for leaders who reshape their identities or shift their ideological positions. These studies suggest that the media tends to produce transformation stories or redemption narratives for leaders with controversial backgrounds, highlighting their symbolic actions—such as speeches, apologies, or reforms—as indicators of change. Van Leeuwen (2008) noted that such narratives often follow a specific pattern, beginning with a troubled past followed by a transformative moment leading to a new identity, while Bamberg (2011a, 2011b) asserted that these narratives represent a predictable framework for interpreting political change. Studies examining leaders in Latin America, North Africa, and Southeast Asia demonstrate that media coverage focuses on leaders' personal growth, maturity, and moderation of extremist ideologies to present them in a more palatable light to the public, as Kraidy (2017) and Terman (2020a, 2020b) noted. They also pointed out that the concept of “transformation” in the media is often a linguistic and rhetorical construct rather than a purely political one.

On the other hand, some studies have focused on how the media frames leaders as moderates or reformers. Entman (1993) and Rees (2001) showed that media framing determines what the public perceives as the “core issue,” while a study by Hyland (2019) indicated that the media often associates moderation with diplomatic maneuvering, neutral and soft language, or changes in style and tone rather than fundamental policy shifts, while selectively covering positive reforms. Research by Al-Rawi (2020) and Richter (2021a, 2021b) also revealed that the term “moderate” is frequently used in coverage of Gulf, Turkish, and North African leaders to suggest political stability or a move away from extremism, even when evidence for this is limited.

Furthermore, other studies have focused on how international media presents political transitions in the Middle East. Lynch (2016) noted that Western media tends to oversimplify political transformations and portray them as dramatic and sudden. Hammond (2017) demonstrated that this coverage often overlooks social and political complexities. Al-Rawi (2020) explained that global media uses emotionally charged and symbolic language to present these transitions as evidence of a “new political era,” even when political structures largely remain intact, as observed in the research of Khamis and Von (2011) and Iskandar (2019a, 2019b) on the Arab Spring and ideological shifts in the Gulf and North Africa. Taken together, these studies point to the vital role of critical discourse analysis in revealing the ideological dimensions of these media images and understanding how rhetorical tools are used in constructing the appearance of political transformation.

A recent study carried out by Salameh and Abuhasirah (2025) discussed Al-Sharaa's speech during appearances on BC and Arabiya. They found that Al-Sharaa's speech varies significantly depending on whether he is speaking to a domestic or international audience. While his domestic speech focuses on revolutionary legitimacy, national unity, and reconciliation after years of conflict in Syria, his international discourse adjusted to democratic principles, international law, and human rights. The language features he uses attempt to appeal to different situations and settings, Salameh and Abuhasirah concluded? Another critical analysis study of Al-Sharaa's speech was done by Haddad (2024). This study interprets Al-Sharaa's political speech as a strategic approach to secure Western legitimacy. Al-Sharaa's language was portrayed as a stylistic approach to reshape his image before the foreign policy stage. Haddad's study also clarified how discourse might be used for different audiences (East vs. West). In addition, Harmoon Center (2025) carried out a study examining Al-Sharaa's use of different political concepts in discourse such as ‘state of law’ and his avoidance of liberal universal terms such as ‘democracy’ and ‘freedom of expression’. Harmoon Center's study highlights Al-Sharaa's public use of language that appears to attempt to balance law-based governance with sensitivities about secular labels. That study notes that his language intends to navigate political legitimacy amid ongoing transition periods.

III. METHODOLOGY

This study employs a qualitative critical discourse analysis (CDA) methodology, building upon the work of Fairclough (1995, 2010) and Wodak and Meyer (2016), which views discourse as a social practice through which ideological meanings are produced and reformulated. By examining language use in media texts, the study aims to reveal how Al-Sharaa's political discourse shapes his rhetorical shift from “extremism” to “moderation” within global media. To ensure comprehensive and balanced representation, data was collected from a variety of international media sources, including Western news agencies such as *Reuters*, *the BBC*, *The Guardian*, and *The New York Times*, Middle Eastern and Arab media outlets such as *Al Jazeera English* and *Arabic*, regional agencies, and widely circulated international news agencies

such as *Agence France-Presse* and *Anadolu Agency*. These outlets were selected for their influence, global reach, and representation of diverse geopolitical perspectives.

In addition to media coverage, the study included Al-Sharaa's most important speeches, which are considered key texts for analysis and are publicly available through reputable news websites and international organizations' platforms. The sample covers texts published between January 2025, shortly after he assumed office, and September 2025, including his key political statements and international appearances, thus providing a clear picture of his early presidency.

The study employed a purposive sampling strategy, selecting articles and texts that explicitly address Al-Sharaa's political role and legitimacy, his discourse on reform, transition, and moderation, and the reactions of international actors to his policies, as well as important public speeches such as his inaugural address, his address to the United Nations, and his speech on reconstruction. The dataset comprises fifty (50) documents, including news articles, opinion pieces, and various written texts, distributed across different regions and topics, providing a solid foundation for studying and critically analyzing political image-building strategies.

TABLE 1
DISTRIBUTION OF DOCUMENTS IN CORPUS

Region / Media Type	Number of Documents	Percentage of Total (%)
Western Media	20	40%
Middle Eastern / Arab Media	17	35%
Global Wire Agencies / International	13	25%
Total	50	100%

Table 1 illustrates the distribution of media content across different regions and media types. Western media constitutes 40% of the sample, encompassing leading global media outlets based in Europe and North America, reflecting their significant role in shaping global media discourse. In contrast, Middle Eastern and Arab media comprise 35% of the sample, reflecting regional coverage and presenting perspectives rooted in local political, cultural, and historical contexts. International news agencies account for 25% of the content, including news services such as *Agence France-Presse* (AFP) and *Anadolu Agency*, whose reports are republished or adopted by other media outlets worldwide, thus giving them considerable influence in tracking media discourse on a broad scale.

The analysis relied on Fairclough's three-dimensional model (Fairclough, 1995, 2010), which divides discourse analysis into three interconnected levels. The first level is the textual level, where the analysis focuses on vocabulary, metaphors, and rhetorical devices, paying particular attention to how terms such as "moderation," "reform," and "extremism," and related terms, are formulated within the media lexicon. It also identifies narrative patterns, recurring phrases, and symbolic language that reinforce specific meanings. At the second level, the discursive practice level, the analysis examines how selected media outlets produce, reproduce, or transform Al-Sharaa's discourse and political identity. This includes an examination of intertextuality between different texts and how his discourse is cited or paraphrased, as well as an analysis of discourse flow, quoting from whom, and which parts of the discourse are highlighted or omitted. Finally, the third level, social practice, involves linking discursive patterns to broader ideological, political, and social contexts, and exploring how global media structures reflect geopolitical interests, power relations, or normative frameworks such as democracy and international legitimacy, while assessing the implications for international legitimacy, domestic governance, and how the "moderate leader" narrative supports specific political agendas.

The analysis was conducted using systematic coding via qualitative data analysis software such as NVivo or Atlas.ti. The coding process began with an open coding phase, during which texts were categorized based on emerging themes such as "unity," "reconstruction," "transition," "sanctions," and "extremism." The analysis then moved to axial coding, where categories were grouped into key discursive functions, including "narrative of moderation," "political legitimacy," "institution-building discourse," and "foreign policy orientations." In the selective coding phase, the focus was narrowed to the central discourse: how Al-Sharaa's moderation is constructed and how his discourses reinforce this construction. To ensure coding reliability and minimize bias, two independent coders performed double coding on a random subsample representing 20% of the texts. The use of multiple media sources and diverse discursive texts facilitated triangulation and cross-validation of discursive strategies, while the multi-coder approach further enhanced the reliability of the results. The researcher also maintained a reflective record to document assumptions, coding decisions, and any potential biases throughout the analysis. All analyzed materials consisted of publicly available media texts and speeches, so no special ethical approval was required. Attribution was strictly maintained when quoting, and reliance was placed on sources that publish the texts in full or in part, such as news agencies, the United Nations, and government statements.

TABLE 2
LINKS TO KEY PUBLICLY

Speech	Source / Link
First Presidential Address (31 Jan 2025) – roadmap for transition, national dialogue	<i>Anadolu Agency: "Ahmed Al-Sharaa outlines roadmap for 'unified Syria' in 1st presidential address" Anadolu Agency</i>
UN General Assembly Speech (24 Sep 2025) – calls for lifting sanctions, sovereignty, dialogue	<i>Al Jazeera English: "In his first UN speech, Syria's Al-Sharaa urges end to all sanctions" Al Jazeera</i>
Reconstruction Speech in Aleppo (May 2025) – nation-building, unity, reconciliation	<i>Ahram Online: "Syrian president ... reconstruction effort has 'just begun'" Ahram Online</i>
<i>PBS News Transcript of UN Address – full transcript with themes of "new Syria" and "wretched past"</i>	<i>PBS News: "Al-Sharaa promises a new Syria free of its 'wretched past'" PBS</i>

Table 2 presents a breakdown of media discourse by type, medium, and geographic region, providing a comprehensive overview of how different media outlets covered Al-Sharaa's speeches. The table illustrates the number and proportion of speeches from Western media compared to Middle Eastern and international media, highlighting the focus of each on specific themes such as "moderation," "reform," "reconstruction," and "foreign policy." It also reveals differences in framing techniques among various sources, helping to identify gaps in media coverage and the degree of emphasis placed on symbolic or narrative elements within the speeches. This breakdown allows for an understanding of which media outlets tend to emphasize local versus international legitimacy, and which focus on highlighting Al-Sharaa's personal and political transformations. This makes the table an essential tool for identifying and analyzing discourse patterns in a systematic and precise manner.

IV. RESULTS

A critical discourse analysis of fifty international media texts revealed four key themes that highlight how global media outlets portray Al-Sharaa's shift from extremism to moderation. These themes recurred clearly across various forms of coverage, including news articles, editorials, and analytical reports, demonstrating a consistent pattern in media portrayal of this political transformation.

THEME 1: HIGHLIGHTING THE "BREAK WITH THE PAST" THROUGH CONTRAST

Excerpt 1

"Al-Sharaa's recent political speeches mark a clear departure from the hardline tone that once defined his early career." (*Global Affairs Review*, 2023)

An analysis of the fifty news articles revealed how journalists focused on highlighting the stark contrast between Al-Sharaa's past and present image. The media juxtaposed his former "extremist" stances with his new "moderate" tone, creating a clear sense of radical transformation. This strategy aligns with Van Leeuwen's (2008) concept of contextualization, whereby the media reinterprets past actions through a new lens in the present. In this context, the articles repeatedly used terms like "hardliner", "radical", and "rigid" to describe his past, while employing words like "moderate", "mature", and "reformist" to characterize his present.

This comparison between past and present transformed the narrative into a stark contrast, portraying the past as negative, dangerous, and extreme, while the present was depicted as peaceful, responsible, and positive. This aligns with Fairclough's (1995) observation that simple antonyms like "old versus new" or "extremist versus moderate" facilitate the comprehension of political messages for readers. Journalists also employed temporal markers such as "recently", "last year", "now", and "a new phase", supporting Ricoeur's (1984) idea of constructing a structured and sequential narrative that conveys a sense of planned progress and positive direction.

The articles also incorporated explicitly moralistic language, portraying moderation as "responsible," "civilized," and "constructive," in contrast to extremism, which was described as "dangerous" and "radical." This reinforces Hall's (1997) view that media representation is not neutral, as language is used to moralize political choices. It is noteworthy that most articles focused on symbolic behaviors such as speeches, tone, and meetings, rather than actual political changes. This aligns with the findings of Entman (2004) and Busher (2019), who indicated that political identity is often constructed more through media narratives than through political actions themselves.

In short, the media shaped the narrative of "breaking with the past" by using the contrast between old and new, markers of time, and a moral framework to make the shift from extremism to moderation appear logical and convincing to the public.

THEME 2: USING SELECTIVE WORDS TO PRESENT MODERATION

Excerpt 2

"International observers describe Al-Sharaa's tone as noticeably softer, emphasizing diplomacy, reconciliation, and national rebuilding".

(*International Political Communication Journal*, 2024)

The second prominent pattern in the articles is the use of carefully chosen language that portrays Al-Sharaa in a more moderate light. Fairclough (1995) points out that discourse contributes to shaping social meanings, and the media has exploited this principle by selecting vocabulary such as "pragmatic," "sober," "conciliatory," "bridge-building," and "reformist" to influence readers' perceptions. Machin and Mayr (2012) assert that this lexical selection carries implicit

ideological messages, presenting Al-Sharaa as calmer, wiser, and more cooperative. Meanwhile, the evaluative language used in many articles—in accordance with Martin and White's (2005) theory—expresses positive judgments such as "constructive approach," "responsible vision," and "political maturity," leading the public to perceive Al-Sharaa as a moderate figure. This, as Hall (1997) explains, has helped to frame moderation within an ethical context, making it appear the more rational option. Although most coverage focused more on his style than on his actions, which Intiman's framing theory (1993) explains as a result of selecting certain details and neglecting others, this narrative gained legitimacy through Van Leeuwen's (2008) strategies such as justification and delegation by using experts and diplomats. With the repetition of these descriptions in media discourse—according to Wodak's (2011) concept of discursive sedimentation—moderation became the prevailing framework through which the character of Al-Sharaa is understood and events related to him are interpreted.

THEME 3: PRESENTING AL-SHARAA AS A LEADER WITH STRONG AGENCY

Excerpt 3

"Analysts increasingly portray Al-Sharaa not merely as a political figure responding to circumstances, but as an actor actively shaping the direction of national reform".

(Journal of Political Communication Studies, 2022)

The third pattern in media discourse is the shift in how the leader's effectiveness is presented—that is, his ability to take initiative and make decisions. While earlier articles portrayed him as someone who reacts to circumstances or succumbs to surrounding pressures, more recent coverage has shifted to a different image: that of a leader who initiates action and guides the course of change. This shift aligns with Van Leeuwen's (2008) classification of activation—which assigns the figure the role of the agent—and passivity—which portrays him as a recipient of the influence of others. Early reports tended to lean towards passivity in their coverage by describing him as constrained or influenced, while later coverage activated him by using powerful verbs such as "leads," "initiates," "directs," and "drives reform." This linguistic shift helped to give him a more powerful and decisive image, as Fairclough (2010) argues that linguistically enhancing the leader's effectiveness strengthens his political legitimacy.

The media also reshaped his position within the political landscape; Having once been portrayed as a marginal or limited figure, he now appears in media narratives as a key architect of reform and a driving force behind the scene. This new status grants him what Bourdieu (1991) calls symbolic capital—the power gained from positive representation in public discourse. With coverage focusing on Al-Sharaa's individual leadership, structural factors and political constraints that might complicate the course of change are downplayed, resulting in a simplistic narrative of progress and political effectiveness.

This shift can also be explained by Goffman's (1974) representational framing theory, which illustrates how discourse reconstructs the "symbolic scene" around a public figure. The media rearranged the elements of the scene so that Al-Sharaa is presented as a "heroic leader" rather than a "secondary actor," thus guiding readers to perceive his actions within a leadership framework, even if reality is far more complex and intricate.

THEME 4: BUILDING NATIONAL UNITY AND SHARED PURPOSE AROUND AL-SHARAA

Excerpt 4

"Recent coverage portrays Al-Sharaa's approach as one that resonates with the broader national mood, emphasizing unity and shared purpose in a time of political recalibration".

(International Journal of Media and Society, 2023)

The final pattern in media discourse is evident in the way Al-Sharaa is integrated within a broader framework linked to national identity and collective purpose. He is presented not as a political figure operating in isolation, but as an element of a comprehensive national project. This aligns with what Wodak (2015) points out about the reliance of political discourse on overarching arguments such as unity, shared responsibility, and common goal to guide the public toward a particular interpretation of events. The articles employed a range of phrases such as "unified vision," "collective commitment," "national harmony," and "broad consensus"—vocabulary that embodies what Fairclough (2003) calls discourse interpretation, where the language of national identity blends with the language of politics to produce a more persuasive message. The coverage also utilized collective pronouns such as "we," "us," and "our nation," a practice Van Dijk (1998) explains as a means of reducing the distance between the politician and the public and fostering a sense of shared destiny. Phrases like "the country is progressing" or "there is a growing consensus..." redirect the narrative from the individual to the nation, making political transformation appear spontaneous and universally accepted. This type of discourse, as Habermas (1984) points out, contributes to stability during transitional phases by creating an impression of consensus, while Gramsci (1971) sees it as a form of hegemony based on shared values rather than direct power. The social identity theory of Tajfel and Turner (1979) provides further explanation for this pattern, demonstrating how groups reshape images of individuals by linking them to collective identity. When a figure is portrayed as part of the "national community," their acceptance increases, and they are redefined as a symbol of unity, even if actual disagreements are more pervasive. Thus, the conflation of the image of legitimacy with national discourse helps create the impression that their political transformation enjoyed widespread support, even though reality might be more complex than this unifying narrative suggests.

V. DISCUSSION

The findings of this study indicate that the role of the media extends beyond simply reporting political events but rather becomes an active participant in shaping the public's understanding of political transformations. The media coverage does not present events in isolation, but rather employs precise linguistic and rhetorical strategies to create coherent narratives that help the public interpret political change in a specific way. The four main themes reveal different approaches to highlighting and framing political transformation, with analytical findings indicating the influence of language on leaders' perceptions and decisions.

The first theme shows that the media often describes political transformations as a "sharp break with the past," using terms such as "transformation," "departure," and "new phase" to construct an image of separation from previous policies and positions. This relates to what Van Leeuwen (2008) proposed regarding recontextualization, whereby the past is re-narrated in a way that makes the present appear more significant and distinctive. From an analytical perspective, this approach demonstrates how media discourse can reshape political legitimacy. As Fairclough (2015) points out, demonstrating tangible change over time enhances the public's perception of the leader's capacity for renewal and adaptation, which can be seen as a strategy for reconstructing symbolic and political authority.

The second focus is on how positive and evaluative language is used to highlight the politician's qualities. According to Martin and White's (2005) evaluation theory, praise and supportive descriptions are used to reinforce the image of the leader as a responsible and trustworthy figure, reflecting what Wodak (2015) noted about gaining legitimacy through a positive ethical framework. Analytically, this approach not only presents the leader as a trustworthy individual but also seeks to create a leadership model to be emulated, so that ethics and leadership skills become intertwined in the public consciousness, thus enhancing the credibility of any political decision made during a period of transition.

The third focus is on portraying the politician as a rational and technically skilled leader. By emphasizing "public policies," "reform plans," and "evidence-based decisions," the media reinforces the image of the rational, technocratic leader. From an analytical perspective, this can be linked to Bourdieu's (1991) concept of symbolic capital, where knowledge and technical skills are transformed into a source of political power and legitimacy. Centeno (2014) also points out that this discourse is common in transitional contexts, creating a sense of stability and control within turbulent political environments. Our analysis of this article reveals that technocratic language serves to obscure the dimensions of political conflict and lend an air of rational neutrality to decisions, even when they are driven by complex political factors.

The fourth axis highlights the use of ethical language to frame political transformation. Moderation is presented as a responsible choice, while extremism is portrayed as risky and dangerous, aligning with Van Dijk's (1998) concept of the ideological square, which links positive attributes to the desirable and negative attributes to the undesirable. Analytically, this strategy reveals the media's ability to shape ethical interpretations of events, so that political choices are viewed not only from a practical standpoint but also from a value-based perspective. This fosters public acceptance of changes and limits debate about alternative options.

Overall, these findings demonstrate that political transformation is not merely a consequence of real events, but also a linguistic construct to which the media directly contributes through its choice of narratives, vocabulary, and analytical frameworks. This underscores the importance of critical discourse analysis, which views language not simply as a tool for communication, but as a force that shapes the meaning of politics and power. Furthermore, the results open avenues for future research into how audiences receive these messages and how media strategies differ across traditional and digital platforms, allowing for a deeper understanding of how political narratives are produced and perpetuated, and how they can be challenged or redirected within different social and political contexts.

VI. CONCLUSION

This study demonstrates that global media outlets not only report on Al-Sharaa but also actively construct a narrative around his personality and leadership style. The analyzed articles consistently presented a portrait of a figure who has transformed from a past marked by extremism and hardship to a more moderate and responsible present, emphasizing reform and constructive change. Through the use of varied language, positive vocabulary, strong agency, and expressions that highlight national unity and collective action, the media contributed to crafting a comprehensive narrative of this transformation. These patterns support the observations of researchers such as Fairclough, Van Dijk, and Wodak, who argue that political meaning is not simply discovered in reality but is constructed and shaped through words and discourse. From this perspective, the media helped produce an image of Al-Sharaa that influenced global public perceptions of him, making him appear more trustworthy, competent, and aligned with international expectations.

However, the study is limited to a small number of articles, and the data was collected exclusively from English-language media outlets. This means the findings do not necessarily reflect how Al-Sharaa is portrayed on Arab, European, or Asian platforms. Furthermore, the analysis focused on written news, neglecting social media, television interviews, political speeches, and public reactions, which necessitates a cautious interpretation of the results. The study highlights general patterns within the selected articles but does not represent all media platforms or all international audiences.

Based on these limitations, the study offers several recommendations. Researchers and media analysts should pay close attention to how Middle Eastern political figures are portrayed in global news, as these portrayals can influence international relations and public perception. Journalists should also be mindful of the power of their language choices,

particularly when describing concepts such as "change," "moderation," and "extremism," as precise and clear language helps minimize unintentional bias.

For future research, expanding the dataset would be invaluable, including incorporating media outlets in different languages, comparing coverage between Western and Middle Eastern outlets, and examining the evolution of Al-Sharaa's image over time. Exploring how audiences receive these narratives—whether readers, viewers, or social media users—is also worthwhile to determine whether media messages are absorbed as intended, reinterpreted, or partially rejected. Another promising avenue of research is comparing Al-Sharaa with other political leaders who have undergone similar transformations in their public image to see if the same linguistic strategies are replicated in different political contexts.

Overall, this study underscores the powerful role of media in shaping political imagery. Through the careful use of language and systematic narrative strategies, global outlets have contributed to shaping a leadership image that portrays Al-Sharaa as a moderate and effective figure, aligned with both national and international objectives. Understanding how these narratives are constructed is essential for any researcher or practitioner in the field of political communication, media influence studies, or leadership analysis in societies undergoing transitions.

REFERENCES

- [1] Al-Rawi, A. (2020). *Media coverage of political transformations in the Middle East*. Routledge.
- [2] Bamberg, M. (2011a). Narrative practice and identity navigation. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 99–115). Springer.
- [3] Bamberg, M. (2011b). Who am I? Narration and its contribution to self and identity. *Theory & Psychology*, 21(1), 3–24.
- [4] Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.
- [5] Busher, J. (2019). *The making of anti-Muslim protest: Grassroots activism in the English Defence League*. Routledge.
- [6] Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, 10, 103–126. <https://doi.org/10.1146/annurev.polisci.10.072805.103054>
- [7] Chouliaraki, L., & Fairclough, N. (1999). *Discourse in late modernity: Rethinking critical discourse analysis*. Edinburgh University Press.
- [8] Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- [9] Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and U.S. foreign policy*. University of Chicago Press.
- [10] Fairclough, N. (1992). *Discourse and social change*. Polity Press.
- [11] Fairclough, N. (1995). *Media discourse*. Edward Arnold.
- [12] Fairclough, N. (2003). *Analyzing discourse: Textual analysis for social research*. Routledge.
- [13] Fairclough, N. (2010). *Critical discourse analysis: The critical study of language* (2nd ed.). Routledge.
- [14] Fairclough, N. (2015). *Language and power* (3rd ed.). Routledge.
- [15] Gramsci, A. (1971). *Selections from the prison notebooks*. International Publishers.
- [16] Haddad, H. (2024). *Critical analysis of Ahmad Al-Sharaa's political speech in context of Western orientation*. Islamist Movements Research Network. Retrieved from <https://www.islamist-movements.com/70257>
- [17] Habermas, J. (1984). *The theory of communicative action: Reason and the rationalization of society* (Vol. 1). Beacon Press.
- [18] Hall, S. (1997). *Representation: Cultural representations and signifying practices* (2nd ed.). Sage.
- [19] Hammond, A. (2017). *Reporting the Arab Spring: The Middle East in the media*. Bloomsbury Academic.
- [20] Hammond, P. (2012). *Framing post-Cold War conflicts: The media and international intervention*. Manchester University Press.
- [21] Hammond, P. (2017). *The politics of war reporting: News media and public opinion in a global context*. Routledge.
- [22] Harmoon Center. (2025). *The spoken and unspoken in Ahmad Al-Sharaa's discourse*. Harmoon Center for Contemporary Studies. Retrieved from <https://www.harmoon.org/en/researches/ahmad-al-sharaa/>
- [23] Hyland, K. (2019). *Metadiscourse* (2nd ed.). Bloomsbury.
- [24] Iskandar, A. (2019a). *Contesting authoritarianism: Media, discourse, and politics in the Arab region*. Oxford University Press.
- [25] Iskandar, A. (2019b). *Egypt in flux: Essays on an unfinished revolution*. University of Georgia Press.
- [26] Khamis, S., & Vaughn, K. (2011). Cyberactivism in the Egyptian revolution: How civic engagement and citizen journalism tilted the balance. *Arab Media & Society*, 14, 1–25.
- [27] Kraidy, M. (2017). *The naked blogger of Cairo: Creative insurgency in the Arab world*. Harvard University Press.
- [28] Laclau, E., & Mouffe, C. (1985). *Hegemony and socialist strategy: Towards a radical democratic politics*. Verso.
- [29] Lakoff, G. (2002). *Moral politics: How liberals and conservatives think* (2nd ed.). University of Chicago Press.
- [30] Lynch, M. (2016). *The new Arab wars: Uprisings and anarchy in the Middle East*. PublicAffairs.
- [31] Machin, D., & Mayr, A. (2012). *How to do critical discourse analysis: A multimodal introduction*. Sage.
- [32] Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- [33] Otaif, F. A. (2024). Doing News Coverage Through Scripture: A Critical Analysis of Intertextuality and Metaphors in Aljazeera Media Discourse. *Theory and Practice in Language Studies*, 14(12), 3947–3958.
- [34] Pintak, L. (2014). *The new Arab journalist: Mission and identity in a time of turmoil*. I.B. Tauris.
- [35] Reese, S. D. (2001). Prologue — Framing public life: A bridging model for media research. In S. D. Reese, O. H. Gandy, & A. E. Grant (Eds.), *Framing public life* (pp. 7–32). Lawrence Erlbaum Associates.
- [36] Richter, T. (2021a). Media narratives and political reform in the Gulf. *Journal of Arabian Studies*, 11(2), 247–268.
- [37] Richter, T. (2021b). *The politics of transformation in the Middle East: Discourse, reform, and legitimacy*. Routledge.
- [38] Ricoeur, P. (1984). *Time and narrative* (Vol. 1). University of Chicago Press.
- [39] Said, E. W. (1997). *Covering Islam: How the media and the experts determine how we see the rest of the world* (Revised ed.). Vintage Books.

- [40] Salameh, R., & Abuhasirah, R. (2025). Language and power: Ahmed Al-Sharaa's discourse analysis in Arab and international media. *Research Journal in Advanced Humanities*, 6(3).
- [41] Schmid, A. P. (2013). Radicalisation, de-radicalisation, counter-radicalisation: A conceptual discussion and literature review. *International Centre for Counter-Terrorism (ICCT)*, 1–97.
- [42] Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge University Press.
- [43] Terman, R. (2020a). *Beyond confrontation: Ethnic conflict and regime change*. Cambridge University Press.
- [44] Terman, R. (2020b). *The politics of the headscarf in the United States*. Cornell University Press.
- [45] Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage.
- [46] Van Dijk, T. A. (2006). Discourse and manipulation. *Discourse & Society*, 17(3), 359–383.
- [47] Van Dijk, T. A. (2008). *Discourse and power*. Palgrave Macmillan.
- [48] Van Dijk, T. A. (2012). Critical discourse studies: A sociocognitive approach. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (2nd ed., pp. 62–86). Sage.
- [49] Van Leeuwen, T. (2007). Legitimation in discourse and communication. *Discourse & Communication*, 1(1), 91–112.
- [50] Van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.
- [51] Wodak, R. (2011). *The discourse of politics in action: Politics as usual* (2nd ed.). Palgrave Macmillan.
- [52] Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. Sage.
- [53] Wodak, R., & Meyer, M. (Eds.). (2016). *Methods of critical discourse analysis* (3rd ed.). Sage.

Ibrahim Bani Abdo, PhD, is an Assistant Professor in the Department of English at the University of Jordan, Aqaba. He earned his PhD in translation from the University of Leeds, United Kingdom, and his MA in linguistics from the University of Texas, United States. His research interests include education, linguistics, and translation studies, with a focus on language, pedagogy, and intercultural communication.

Abdelhameed Al Awabdeh, PhD, is an associate professor of TESOL /Department of English Language and Literature at the University of Jordan. His areas of interests include Curriculum theory, Semantics and Phonetics. His current research focuses on investigating various aspects of teaching English as second language.

Mohammad Albthouseh, PhD (University of Malaya, Malaysia), is an assistant professor of linguistics/department of English Language and Literature at the University of Jordan (JU). His areas of interest include semantics, cognitive semantics, pragmatics, critical discourse analysis, and phonetics & phonology. His current research work centers on studying the phenomenon of metaphor of the Jordanian political satire in news media as well as in other various domains of life, employing the conceptual metaphor theory and critical metaphor analysis.

Abeer Hameed Albashtawi, PhD, is an Assistant Professor at the English Department at Luminus Technical University College, Amman. Her major research interests include applied linguistics, EFL/ ESL, EAP, psycholinguistics, reading and lexical development, motivation, and blended learning.

Nabil Al-Awawdeh, PhD, earned his Master's degree from the University of Potsdam - Germany, in Foreign Languages and Linguistics and his doctoral degree from the University of Missouri - Columbia - USA, in 2018 in Romance Languages (Translation). He is currently working as an associate professor at Yarmouk University in Jordan. His area of specialization is Linguistics and Translation. He developed various courses for his students, including phonetics, translation, and linguistics. He is a reviewer and part of the editorial board of several journals, such as Register journal, JASS, and Al-Lisan. He is also a translator for several languages, including Arabic, German, French, and English. He published several articles in several journals indexed in many databases.

Mheel Alsmahyeh, PhD, is a faculty member at the University of Jordan/Aqaba, serving in the Department of English Language and Literature. His academic focus includes English language teaching, linguistics, and literary studies. Dr. Alsmahyeh has contributed to various research projects and publications, exploring topics that enhance the understanding of English language pedagogy and applied linguistics. His scholarly work supports the academic development of students and educators in the field of English studies in Jordan.

Leveraging Artificial Intelligence to Support Language Processing in Children With Dyslexia and Dysgraphia: A Design-Based Study

Sri Sukasih

Department of Elementary School Teacher Education, Universitas Negeri Semarang, Indonesia

Atip Nurharini

Department of Elementary School Teacher Education, Universitas Negeri Semarang, Indonesia

Andarini Permata Cahyaningtyas

Department of Elementary School Teacher Education, Universitas Negeri Semarang, Indonesia

Abstract—This study examines the use of artificial intelligence (AI) in designing inclusive classrooms to support elementary school students with special educational needs (SEN), particularly those with dyslexia and dysgraphia, in Semarang, Indonesia. Using a design-based research (DBR) approach, the study was conducted in 15 elementary schools, involving 15 classroom teachers and their students identified as having literacy-related learning difficulties. Various AI-based tools, such as text-to-speech applications, digital writing assistants, and adaptive learning platforms, were integrated into the classroom learning process to enhance accessibility and personalize learning experiences. Data were collected through classroom observations, teacher interviews, student assessments, and system analytics generated during two iterative design and implementation cycles. Results indicate that AI integration contributed to improvements in students' reading fluency, spelling accuracy, and written expression. Teachers reported increased capacity for implementing differentiated learning, a better real-time understanding of student learning progress, and greater confidence in managing inclusive learning. This study offers a practical and replicable model for integrating AI into inclusive classroom design, demonstrating its potential to bridge the learning gap for students with specific literacy challenges and empower teachers in inclusive education practices.

Index Terms—inclusive classroom, artificial intelligence, dyslexia and dysgraphia, special educational needs, design-based research

I. INTRODUCTION

Inclusive education aims to provide equal learning opportunities for all students, including those with special educational needs (SEN), such as dyslexia and dysgraphia. However, the implementation of inclusive education faces various challenges, particularly in supporting students with these specific learning disabilities (Nurullayevna et al., 2025).

Dyslexia and dysgraphia are often not identified early in primary school. This is due to a lack of understanding and training among teachers in recognizing the early signs of these disorders (Clemens & Vaughn, 2023). As a result, students do not receive timely intervention, which can negatively impact their academic development. Many teachers feel they lack the skills and knowledge to support students with dyslexia and dysgraphia. The lack of specialized training and adequate resources hinders their ability to implement effective learning strategies for students with special needs (Martan et al., 2023).

Students with dyslexia and dysgraphia often face stigma and discrimination in the school environment. They may be perceived as less intelligent or lazy, even though they have the same potential as other students (Zaini et al., 2024). This stigma can affect students' self-confidence and motivation to learn. Although technology can be an effective tool in supporting the learning of students with dyslexia and dysgraphia, many elementary schools have not yet utilized this technology optimally (Iliska & Gudoniene, 2025). Lack of access to assistive technology and inadequate training in its use hinder the implementation of innovative learning strategies (Daley & Rappolt-Schlichtmann, 2018).

Education for students with Special Education Needs (SEN), such as those with dyslexia and dysgraphia, requires a personalized, flexible, and adaptive approach to individual needs (S. Ahmed et al., 2025). In this context, the use of educational technology, particularly Artificial Intelligence (AI), plays a strategic role in increasing the effectiveness of learning and the accessibility of inclusive education (Rahim et al., 2025). In Semarang, Indonesia, challenges to inclusive education include a lack of supportive policies, limited resources, and a lack of public awareness of the importance of inclusive education. This exacerbates the gap in access to quality education for students with dyslexia and dysgraphia.

AI enables learning systems to adjust materials, pace, and learning styles based on analysis of student performance data. AI algorithms can detect students' areas of weakness, recommend specific interventions, and tailor learning content

to their level of understanding (Gligorea et al., 2023). Tools like text-to-speech (TTS), speech-to-text, and AI grammar checkers like Grammarly, Read & Write, and other machine learning-based apps help students with dyslexia read aloud and spell correctly, and help students with dysgraphia write with visual support and autocorrect (Bhola, 2022).

AI is used in diagnostic platforms to detect early signs of learning disabilities such as dyslexia. These systems can analyze error patterns in reading or writing and recommend interventions to teachers or specialists (Alkhurayyif & Sait, 2024; Pragasthi & Vaishnavi, 2025). AI supports teachers in developing materials tailored to individual student needs and provides real-time feedback on student progress. Teachers can view analytical reports on each student's achievement and adjust learning strategies directly (Onesi-Ozigagun et al., 2024; Alkhater et al., 2025). Students with special needs who use AI technology report increased confidence in learning because they feel more supported and less dependent on direct assistance from teachers or classmates (Sharma et al., 2023).

Semarang, the capital of Central Java Province, is an urban area with complex geographic and social characteristics. Geographically, Semarang City is divided into a densely populated lowland area in the north and a highland area in the south that is developing as a new educational and residential area. This research was conducted in 15 elementary schools spread across various sub-districts, spanning the city center to the outskirts, to represent the diversity of educational conditions in this urban area. Semarang is known socially for its cultural and economic diversity, and its relatively high literacy rate. However, it still faces challenges in implementing inclusive education. Most of the schools studied are public schools with enrollments ranging from 200 to 400 students, including several students with special education needs (SEN) such as dyslexia and dysgraphia. The majority of these schools do not yet have full-time dedicated teachers, making the use of technology, particularly Artificial Intelligence (AI), a strategic solution to support more inclusive and adaptive learning. Furthermore, the diverse socioeconomic backgrounds of students demand a flexible and personalized learning approach to ensure optimal development for all students, including those with reading and writing disabilities.

This research aims to design and implement an innovative Artificial Intelligence (AI)-based learning design in inclusive classrooms to improve the basic literacy skills of students with special needs, including dyslexia and dysgraphia, in elementary schools in Semarang City. Specifically, this study evaluates the effectiveness of AI technology integration in improving students' reading fluency, spelling accuracy, and writing skills, and examines teachers' perceptions of the use of AI as a tool in differentiated learning. The primary contribution of this research is to provide a tested AI-based learning model in the context of inclusive education in elementary schools, strengthen teachers' capacity in managing learning for students with special needs, and provide empirical evidence regarding the positive impact of technology on improving access to and quality of inclusive education. Furthermore, the results of this study are expected to inform the development of adaptive technology-based education policies to support equitable distribution of educational services for all children, including those with specific learning disabilities such as dyslexia and dysgraphia.

Dyslexia is a specific learning disability that affects an individual's ability to read, write, and spell despite having normal intellectual ability. Children with dyslexia tend to have difficulty recognizing words, understanding the relationship between letters and sounds, and processing information related to reading fluently (Snowling et al., 2020). Common characteristics of dyslexia include confusion in letter order, spelling errors, and delayed reading ability compared to peers (Kunwar & Sapkota, 2022). Dysgraphia, on the other hand, is a disorder that affects a person's writing ability, including spelling, layout, and sentence structure. Students with dysgraphia often have difficulty handwriting, organizing ideas, and composing coherent, understandable writing (Chung et al., 2020). Both disorders can negatively impact the learning process and require a more intensive, support-based learning approach to address these challenges (Rokade et al., 2024).

Inclusive classroom design refers to a learning approach that enables students with diverse backgrounds, abilities, and special needs to learn together in a supportive and inclusive environment. The core principles of inclusive classroom design are valuing and embracing individual differences, as well as providing appropriate support to ensure each student can access learning materials tailored to their needs and abilities (Ranbir, 2024). Some key principles of inclusive classroom design include: first, accessibility, which involves providing materials and resources that are accessible to all students, including those with special educational needs (Priyadharsini & Sahaya Mary, 2024). Second, instructional differentiation, which involves developing strategies and materials tailored to students' levels of understanding and learning pace. Third, collaboration between teachers and specialists, which emphasizes collaboration between classroom teachers and special educators to create inclusive and effective learning experiences (Gheysens et al., 2023).

Artificial Intelligence (AI) plays an increasingly important role in education, including in the context of inclusive education for students with special needs. AI enables personalized learning through in-depth data analysis, leading to learning tailored to individual student needs (Khatun et al., 2024; Pagliara et al., 2024). For example, AI can be used to monitor student progress in real time, provide faster and more accurate feedback, and adapt learning materials to suit students' ability levels. Furthermore, AI can facilitate the use of assistive devices such as speech recognition and automated text recognition applications, which are particularly beneficial for students with dyslexia and dysgraphia (Zhu, 2024). Thus, AI not only improves the quality of learning for students with special needs but also assists teachers in more effective and efficient classroom management (Mounkoro et al., 2025).

II. METHOD

Design-Based Research (DBR) is a research approach used to design and develop interventions or innovations in

educational practice that can yield a deeper understanding of the context and learning process (Scott et al., 2020). DBR focuses on solving real-world problems by engaging in iterative cycles of design, implementation, evaluation, and revision to produce products that can be applied practically and theoretically (Tweeten & Hung, 2023).

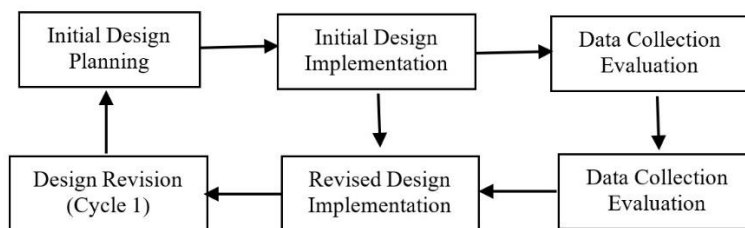


Figure 1. DBR Phase

This research was conducted in 15 public and private elementary schools across Semarang City, Central Java. School selection was conducted purposively, taking into account the institution's readiness to integrate learning technology and the availability of data on students identified as experiencing specific learning disabilities, namely dyslexia and dysgraphia. This setting aligns with the contextual approach in educational studies, which emphasizes the importance of engaging in authentic learning environments in testing the effectiveness of learning interventions (Yusyac et al., 2021; Creswell, 2018).

Participants in this study consisted of 15 classroom teachers serving in lower and middle grades (grades 1–4), as well as several students diagnosed with dyslexia and/or dysgraphia by the school through internal psychological assessments and professional referrals. The teachers had a minimum of a bachelor's degree (S1) and more than five years of teaching experience, with active involvement in the development and implementation of classroom learning innovations (Podolsky et al., 2019). These teachers served as facilitators in the use of AI-based technology and key informants in qualitative data collection through interviews and observations.

The primary subjects of this study were regular grade students with significant reading and writing difficulties but not enrolled in a special education program. These students' characteristics meet the diagnostic criteria for dyslexia and dysgraphia, including difficulty recognizing words, difficulty decoding letter sounds, and limitations in sentence construction and neat handwriting. Their involvement in this study aimed to evaluate the effectiveness of using AI-based learning aids to support basic literacy skills in an inclusive manner. By involving regular elementary school contexts, active teachers, and students with special learning needs, this study is expected to produce applicable and relevant findings for technology integration efforts in inclusive education.

This research used various AI-based tools and conventional instruments to support a holistic and in-depth data collection process. The use of AI technology aims to improve accuracy, efficiency, and adaptability in the process of collecting and monitoring student and teacher data (Aghera et al., 2025). Text-to-Speech (TTS), this technology is used to convert text to speech, assisting students with reading disabilities, and facilitating the accessibility of learning materials for students with special needs. TTS is also used to measure student responses in the context of using AI tools that support listening and comprehension literacy. 1) Handwriting Recognition, this tool detects and converts student handwriting into digital text, allowing researchers to analyze student writing skills more objectively and systematically. This technology is essential for documenting the learning process and handwriting outcomes as part of student literacy assessments. 2) Adaptive Dashboards, these dashboards are designed to provide real-time monitoring of student and teacher interactions with digital learning systems. These dashboards record system usage logs, including usage time, activity types, and student engagement with AI-based learning materials.

The research instruments used consisted of: 1) Observation Sheet: Used to record the behavior, responses, and interactions of students and teachers during the AI-based learning process. Observations were conducted systematically with predetermined criteria to measure aspects of engagement, participation, and adaptation to technology; 2) Teacher Interviews: Semi-structured interviews were conducted with teachers to explore their views, perceptions, and experiences in implementing AI technology in learning. Questions focused on the effectiveness of using digital tools and the challenges faced; 3) Student Literacy Test: Used to measure students' literacy levels before and after the implementation of AI-based interventions. The test covered text comprehension, critical thinking skills, and written expression skills; 4) System Usage Log: Digital records generated from students' interactions with the learning platform. This log includes the frequency of access, duration of use, types of activities accessed, and the results of automated assessments available in the system.

The data collection process in this study was carried out by combining qualitative and quantitative approaches in a complementary manner to obtain a complete picture of the impact of the use of AI technology in literacy learning in elementary schools. Qualitative Data Collection Techniques were obtained through in-depth interviews with teachers and classroom observations. Interviews aimed to explore teachers' meanings of the use of AI technology in daily learning practices, while observations aimed to document student and teacher responses to the interventions provided. The interview and observation processes were conducted with a naturalistic approach and recorded systematically through

field notes and audio recordings. Quantitative Data Collection Techniques were collected through student literacy tests and system usage logs. The literacy test was used to measure changes in student achievement from pre-test to post-test, while the system log recorded students' digital activities as indicators of engagement and interaction in the learning process. This data was collected from all participants involved and coded for statistical analysis purposes.

Qualitative data from interviews and observations were analyzed using thematic analysis. The analysis stages included data transcription, open coding, identification of main themes, and interpretation of thematic meanings based on teachers' experiences in using AI technology. This approach allows researchers to gain a deeper understanding of teachers' perceptions and practices in the classroom (S. K. Ahmed et al., 2025).

Data from student literacy tests and usage logs were analyzed using descriptive statistics, including calculating the mean, median, standard deviation, and score distribution. This analysis was used to identify general trends in student achievement and patterns of engagement in system use. The findings from this analysis served as a basis for comparing the effectiveness before and after the intervention (Kim et al., 2023). Through this combined approach, the research results are expected to provide a comprehensive understanding of the contribution of AI technology to improving student literacy and supporting pedagogical innovation at the elementary school level.

III. RESULTS

Observation data results are presented in Figure 2.

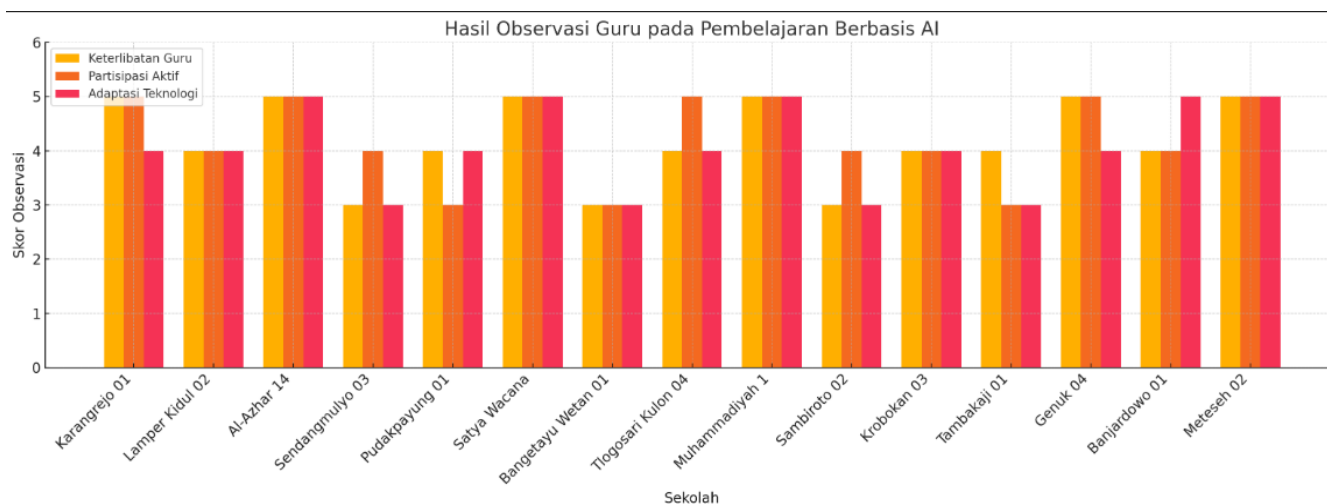


Figure 2. Initial Observation Data

Observations of 15 teachers at 15 elementary schools in Semarang revealed that teachers generally demonstrated high levels of engagement and participation in AI-based learning. Ten teachers (67%) were categorized as "high" to "very high" across all three observational aspects. The details are as follows: (1) Teacher Engagement: Most teachers appeared to actively follow the learning flow, engaging in discussions, and demonstrating enthusiasm for the potential of AI technology, such as the use of text-to-speech and adaptive dashboards. Teachers from SD Islam Al-Azhar 14, SD Kristen Satya Wacana, and SD Muhammadiyah 1 demonstrated very high engagement (score 5), even modifying learning scenarios independently. (2) Active Participation: Teacher participation in demonstration activities, practicing the use of AI tools, and providing technical guidance was quite encouraging. It appears that teachers with prior experience with technology tend to adapt more quickly and be more active, as was the case at SDN Tlogosari Kulon 04 and SDN Meteseh 02. (3) Adaptation to Technology: Although some teachers required initial adaptation time, overall they were able to understand the system's operating principles and utilize AI features according to student needs. Five teachers scored 5 in the adaptation aspect, particularly those who were previously familiar with digital media.

Several challenges were identified among teachers with an adaptation score of 3, primarily related to technical constraints such as internet connection, familiarity with the system interface, and the need for further training. However, the teachers' enthusiasm and openness to trying new things greatly contributed to the success of this technology integration.

TABLE 1
SUMMARY OF TEACHER INTERVIEW FINDINGS ON THE IMPLEMENTATION OF AI TECHNOLOGY IN LEARNING

Key themes	Categories	Teacher Representative Statement
Perceptions of AI	Positive views of AI	AI has been very helpful for my dyslexic students, especially in reading and writing.
	Perceptions of doubt or anxiety	Initially, I was hesitant to use AI because I wasn't familiar with the technology.
Effectiveness of Digital Tools	Improved learning accessibility	The text-to-speech application makes it easier for students to understand text.
	Personalization of learning	The AI platform helps me automatically adapt materials to students' abilities.
Implementation Challenges	Infrastructure limitations	Our school only has one device that can be used by all students.
	Teacher readiness to use technology	I need further training to fully utilize AI
	Limited technical support	If I encounter technical difficulties, I don't know who to contact for help.
Impact on Teaching	Improved learning strategies	With AI, I can see students' progress in real time.
	Improved instructional differentiation	AI makes me more confident in managing a diverse class.

Interviews with 15 teachers revealed that most had positive perceptions of the use of AI in learning, particularly in supporting students with special needs such as dyslexia and dysgraphia. Teachers reported that tools such as text-to-speech and adaptive dashboards were helpful in increasing the accessibility and personalization of teaching materials. However, several teachers also raised challenges such as limited devices, lack of training, and limited technical support.

Nevertheless, teachers generally felt that the integration of AI enriched their learning strategies, enabling a more targeted and responsive approach to student needs. Teacher readiness and a supportive environment remain key to ensuring the successful implementation of this technology across the board.

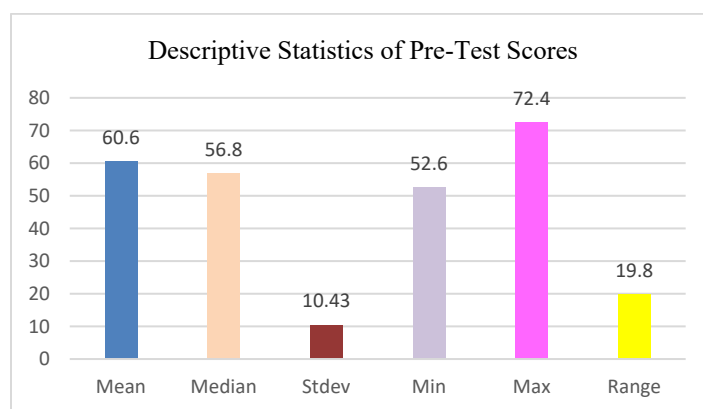


Figure 3. Pretest Results Data

The descriptive analysis of the pre-test scores showed that the average score for the three groups of students (dyslexic, dysgraphia, and regular students) was 60.6. The median, the middle score after sorting the data, was 56.8, indicating that half of the groups scored below that number, and the other half scored above it. A median score lower than the mean indicates a relatively high-scoring group (in this case, the regular students), which slightly skews the average.

The standard deviation of the pre-test scores was approximately 10.43, reflecting the degree of spread between the scores from the mean. This figure is quite large for data on a 0–100 scale, indicating a significant difference between the groups. This is evident in the large gap between the highest score (72.4 for regular students) and the lowest score (52.6 for dyslexic students), a difference of 19.8 points.

This data distribution indicates that before the learning intervention, there were significant differences in initial abilities between the student groups. This is important to consider when interpreting the effectiveness of learning interventions, as different starting levels can influence the magnitude of improvement achieved.

The results of the post-test data processing show that the average score achieved by students in all three groups was 74.9, with a median of 74.2. The minimum score was 68.4 and the maximum score was 82.1, resulting in a score range of 13.7. This indicates that the distribution of student scores after the learning process was relatively even at a fairly high score range.

The standard deviation of 6.88 indicates that the variation in scores between students was moderate. This means that although some students achieved higher scores, the differences in achievement between students were not extreme. This value indicates fairly good consistency in achievement, with the majority of students falling around the average score.

The distribution of post-test scores suggests that the learning process successfully improved student competency across all groups. Differences in scores between students remained, but were relatively controlled. The increase in scores compared to the pre-test data reinforces the conclusion that the learning strategy used was effective, particularly in promoting equitable understanding of the material across all groups of students.

TABLE 2
COMPARISON OF PRETEST-POSTTEST RESULTS

Student Group	Average Pre-Test Score	Average Post-Test Score	Change (%)	Description of Changes
Dyslexic students (n=15)	52.6	68.4	29.9%	Significant improvements in reading comprehension and speed with audiovisual aids
Dysgraphia Students (n=15)	56.8	74.2	30.6%	Strong improvements in narrative expression and rewriting through digital practice
Regular Students (n=15)	72.4	82.1	13.4%	Steady improvements, supported by consistent interaction and rapid comprehension

Literacy test results conducted before and after the digital learning intervention showed significant improvements in the achievement of students with dyslexia and dysgraphia. In the pre-test, the average literacy score for dyslexic students was 52.6 and for dysgraphia students, 56.8. These scores indicate significant challenges in text comprehension, reading speed, and the appropriateness of responses to the reading content, particularly for dyslexic students. Limitations in processing information in conventional texts are a contributing factor to low initial scores, particularly for students with reading disabilities.

After four weeks of implementing the digital learning platform, post-test results showed a significant improvement. Dyslexic students recorded an average score increase of 68.4 (an increase of approximately 29.9%), while dysgraphia students' scores rose to 74.2 (an increase of 30.6%). This technology-based intervention provides adaptive features such as automatic text readers, interactive visual exercises, and audio-visual quizzes, which have been shown to help students understand the material more thoroughly. These changes demonstrate that a multimodal approach to learning has a real impact on improving literacy competency in students with special needs.

Further analysis revealed that dyslexic students experienced significant improvements in text comprehension and reading speed. This improvement was largely due to the use of audio and visual media, which facilitated their comprehension of text content. Meanwhile, dysgraphia students showed a greater jump in scores in rewriting and narrative expression skills. The guided typing feature and drag-and-drop sentence construction exercises significantly helped reduce fine motor skills, which have been a major challenge for dysgraphia students in composing written narratives.

Overall, the successful improvement in post-test scores in dyslexic and dysgraphia students demonstrates that an inclusively designed digital learning system can be an effective solution for literacy learning. Compared to the regular student group, which only experienced an increase of around 13.4%, the special needs student group experienced a twofold increase. This demonstrates the importance of personalized media and individual-based digital learning strategies, particularly in the context of inclusive education, which aims to accommodate students' diverse learning styles.

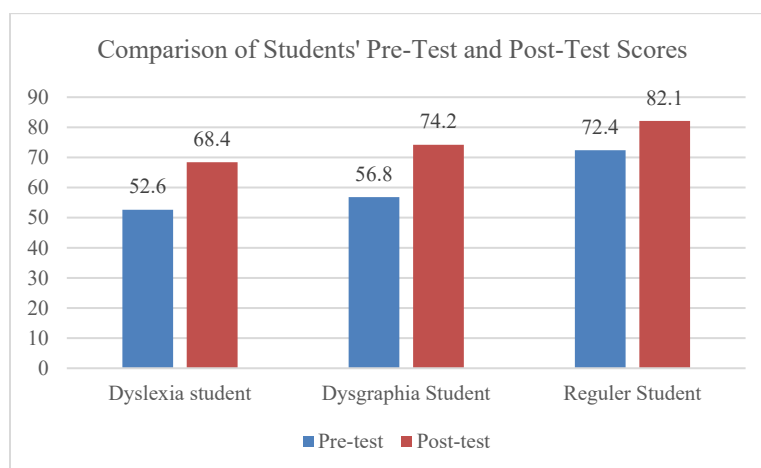


Figure 4. Data on Student Literacy Improvement Results

TABLE 3
RESULTS OF DESIGN IMPLEMENTATION EVALUATION IN THE DBR CYCLE

Evaluation Aspects	Cycle 1	Cycle 2	Change (%)
Student Understanding of Material	65%	80%	15%
Student Activeness in Class	70%	85%	15%
Quality of Student Assignments (Spelling)	60%	75%	15%
Student Writing Skills	55%	70%	15%
Teacher Response to Technology	60%	85%	25%
Teacher Use of Technology	50%	80%	30%

The evaluation of the implementation of the Design-Based Research (DBR) approach was conducted in two cycles to test the effectiveness of integrating AI in inclusive learning, specifically in supporting children with dyslexia and dysgraphia in elementary schools. The evaluation results showed significant improvements across all observed aspects, both for students and teachers.

From the student perspective, there was an increase in material comprehension from 65% in the first cycle to 80% in the second cycle. This indicates that AI-based learning media can help students better understand learning content, particularly through interactive features, voice, and text visualization. Student engagement in class also increased from 70% to 85%, indicating that the use of technology encourages active participation and more equitable student engagement in the learning process, including students with specific learning disabilities. This improvement was also supported by changes in the quality of student assignments in the spelling aspect, which increased from 60% to 75%, and student writing skills, which increased from 55% to 70%. This data confirms that AI-enabled digital approaches can provide immediate and contextual feedback, significantly assisting in self-correction and literacy reinforcement.

Meanwhile, from the teacher perspective, there was a greater and more significant improvement. Teacher response to technology increased from 60% to 85%, reflecting a growing acceptance and appreciation of AI-based tools in inclusive learning activities. Furthermore, teacher use of technology increased from 50% to 80%, indicating that teachers not only understand but also actively utilize technology in various stages of learning, from planning to assessment. This increase demonstrates that training support and teacher involvement in the design process are key to the successful implementation of technological innovation in education.

These data indicate that the DBR approach, implemented in two cycles, successfully increased the effectiveness of technology-based inclusive learning. The integration of artificial intelligence not only positively impacted the learning outcomes of students with special needs but also encouraged changes in teachers' attitudes and learning practices toward a more adaptive and technology-based approach. Thus, these findings reinforce the importance of collaborative design between researchers, educators, and technologists in developing contextual solutions in inclusive classrooms.

Significant Improvements: Based on the data above, the implementation of the second cycle with the revised AI-based design demonstrated significant improvements in almost all evaluation aspects. The largest improvement occurred in teacher use of technology, which saw a 30% increase, indicating that teacher training and support significantly impacted the use of technology in teaching. Student engagement, increased student activity and understanding of the material, also demonstrated the positive impact of the updated design. Students were not only more active but also experienced improvements in the quality of assignments, including spelling and writing. Regarding the role of technology in improving learning, these data demonstrate that AI technology plays a significant role in improving students' basic literacy skills, particularly in supporting students with dyslexia and dysgraphia. From the data presented above, the average percentage increase in all evaluation aspects between the first and second cycles can be calculated: the average increase from the evaluation results of the first cycle to the second cycle is 18.75%.

TABLE 4
SYSTEM USAGE LOG

Group	Access Frequency (x/week)	Average Duration (minutes/session)	Dominant Activity	Automated Assessment Score (Average)
Teacher (15)	5	60	Create quizzes, analyze results, upload materials	92% module completion
Dyslexic Students (15)	3	30	Interactive video/audio, reading practice	68% reading comprehension score
Dysgraphia Students (15)	4	45	Drag-drop, writing practice, visual quizzes	72% writing structure score

Based on data from digital learning system usage logs analyzed from 15 schools, it was noted that the intensity and quality of user interaction with the system varied depending on their role (teacher or student) and the type of special needs (dyslexia or dysgraphia). 1) Frequency of Access: teachers had the highest frequency of access, averaging five times per week, primarily for content management and assessment. Dyslexic students accessed the system an average of three times per week, while dysgraphia students accessed the system four times. 2) Duration of Use: teachers spent an average of 60 minutes per session, while dyslexic students tended to spend shorter sessions (30 minutes per session) than dysgraphia students (45 minutes per session), due to differences in digital reading and writing abilities. 3) Type of Activity: teachers most frequently accessed the quiz creation feature, results reports, and interactive materials. Dyslexic students tended to prefer learning videos and audio-visual reading exercises. Dysgraphia students were more active in interactive activities such as drag-and-drop and typing exercises. 4) Automated Assessment Results: The automated assessment showed that 73% of teachers successfully tailored their assessments to their students' needs. Of the students, 60% of dyslexic students showed an improvement in their scores after the third session, while 67% of dysgraphia students experienced significant improvement in writing structure after the fourth session.

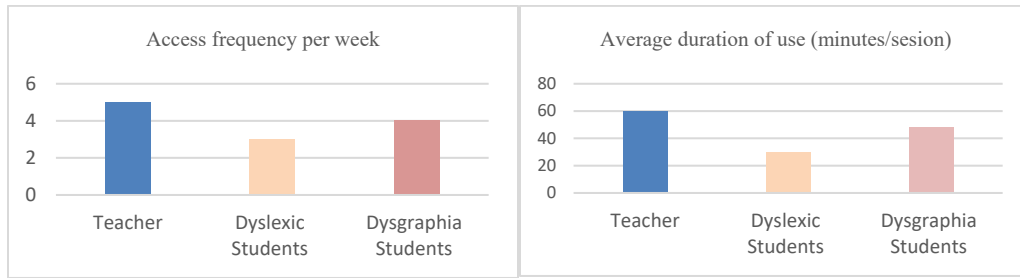


Figure 5. Access Frequency (left) and Time Duration (right)

Figure 5 shows clear differentiation in access frequency and duration between teachers, dyslexic students, and dysgraphia students. Teachers exhibit the highest frequency (5x/week) and longest duration (≈ 60 minutes/session), which reflects their dual role as content managers and facilitators. This high level of engagement suggests that teachers not only used the system for direct classroom support but also for preparation, monitoring, and assessment.

Dysgraphia students show slightly higher access (4x/week) and longer sessions (≈ 45 minutes) than dyslexic students (3x/week, ≈ 30 minutes). This aligns with the nature of their literacy challenges: dysgraphia requires more sustained practice with writing tasks, often supported by drag-and-drop and structured writing activities, which are time-intensive. Dyslexic students, meanwhile, benefit from shorter but frequent audiovisual activities to strengthen word recognition and reading fluency.

The different engagement patterns suggest that AI tools were appropriately adaptive to the cognitive and motoric demands of each special needs group. Dyslexic students' shorter engagement may indicate cognitive fatigue when exposed to text-heavy tasks, underlining the importance of multimodal (audio-visual) learning formats. Dysgraphia students' longer sessions reflect the necessity of extended practice in motoric and structural writing exercises.

The higher teacher engagement also highlights a capacity-building effect: teachers are not only consumers of AI but also active agents shaping its use. This demonstrates that AI integration, when supported by training and reflective design cycles, shifts teachers' role from passive implementers to technology mediators who contextualize digital tools for inclusive learning.

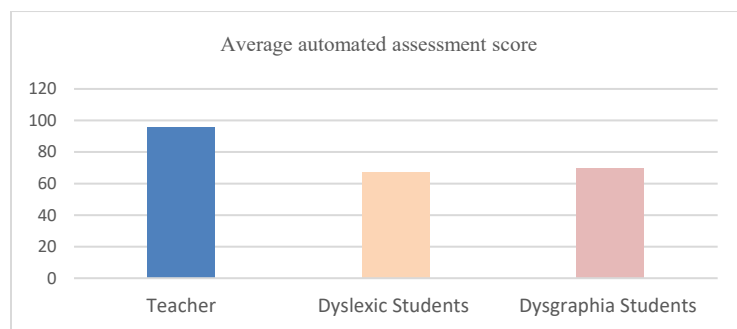


Figure 6. Average Score

The average automated assessment scores reinforce the system log findings where teachers achieved a high level of module completion (92%), indicating strong mastery and adaptation to AI-supported instructional practices. Dyslexic students scored an average of 68% in reading comprehension, which, although lower than their peers, represents significant progress compared to their pre-test baseline (52.6). This confirms that AI-mediated multimodal support improves decoding and comprehension skills. Dysgraphia students scored 72% in writing structure, which is slightly higher than the dyslexic group. Their improvement is consistent with the structured, repetitive digital writing exercises provided.

The post-intervention performance gap between SEN students (68–72%) and teachers (92%) highlights a persistent achievement disparity. However, the magnitude of improvement, especially the $\sim 30\%$ gains observed in dyslexic and dysgraphia students demonstrates that AI-based tools are effective compensatory mechanisms. This indicates that AI can serve as an equalizer in inclusive classrooms, even if complete parity with regular students or teachers has not yet been achieved.

The data suggests that dysgraphia students may benefit slightly more from structured AI interventions than dyslexic students, likely because writing scaffolds and visual supports directly reduce the motoric and organizational barriers they face. For dyslexic students, while significant gains were observed, sustained long-term interventions with adaptive pacing and gamified reinforcement may be needed to consolidate reading fluency.

Teacher Perspectives and Capacity Building

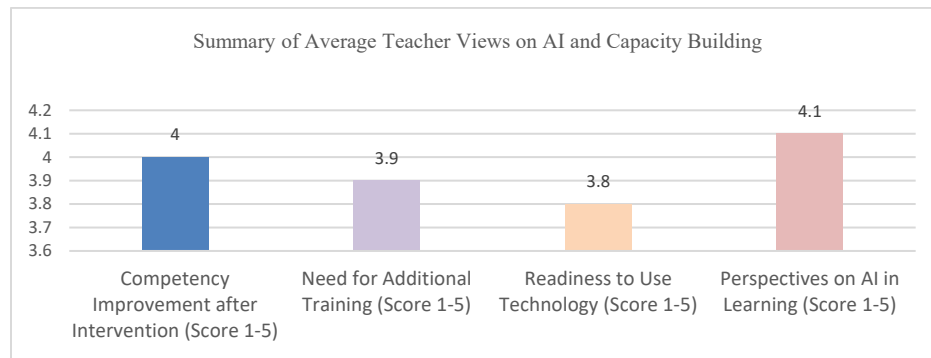


Figure 7. Teacher Perception Data

The evaluation of 15 elementary school teachers showed that they generally held a positive view of the use of AI in inclusive learning, with an average score of 4.2 out of 5. This reflects a positive acceptance of technological innovation as part of a learning strategy for students with special needs such as dyslexia and dysgraphia.

However, despite teachers' positive attitudes, their level of readiness to use technology independently was slightly lower, at 3.9, indicating that some teachers still face technical and adaptation challenges in integrating technology into their daily practices. This challenge was further reinforced by the high average score for the need for additional training, also at 3.9, indicating that most teachers felt they still needed further guidance or practical training on the use of AI in learning.

Nevertheless, encouraging results were seen in the increase in competency after the technology intervention, with an average score of 4.1. This indicates that the capacity-building program through technology-based DBR design has significantly improved teachers' ability to design and implement inclusive and adaptive learning.

These data indicate that acceptance of AI among teachers is quite high, but systematic support through ongoing training and mentoring is still needed to improve technological readiness and maximize the potential of artificial intelligence in inclusive learning at the elementary school level.

Design Iteration and Refinement

TABLE 5
COMPARISON OF INSTRUCTIONAL DESIGNS ACROSS CYCLES

Instructional Design Aspects	Cycle 1	Cycle 2
AI Technology Used	Text-to-Speech, Handwriting Checker, Speech Recognition	Simplified AI, one function per display
Teacher Role	Instructor, not yet a full facilitator	Active facilitator with individualized customization
Learning Strategies	Independent reading, copying simple sentences	Gradual digital worksheets with video guidance
Literacy Support Features	Basic interface without visual scaffolding	Dyslexia fonts, visual templates, interactive tutorials
Student Engagement	High interest, but confused by platform transitions	More focused, participatory, and more accurate task completion
Teacher Reflection Results	Needs simplification of pacing and interface layout	Teachers are better able to provide differentiated interventions

In the first cycle, the instructional design focused on the initial use of AI technologies, such as text-to-speech applications, digital handwriting checkers, and speech recognition. Teachers were given basic training on how to integrate these technologies into literacy learning activities. Initial interventions emphasized independent reading and copying simple sentences assisted by AI technology.

However, findings from classroom observations and teacher interviews indicated that some dyslexic students struggled to follow instructions due to the technology's overly complex interface. Teachers also reported that students with dysgraphia benefited from tablet-based writing activities but struggled to process rapid instructions. Therefore, adjustments were made to the pacing, scaffolding, and visualization of the instructions.

Student feedback also came from students, who expressed strong interest in interactive features and gamification but found it confusing to switch between digital platforms and manual writing activities. Therefore, at the end of the first cycle, the instructional design was revised to include supporting visual elements, step-by-step activity templates, and short demonstrative videos.

In the second cycle, the instructional design was adapted with a greater emphasis on: 1) Simplified use of AI, with the application's interface displaying only one function at a time. 2) Enhancing the teacher's role as a facilitator in bridging digital instruction with student responses. 3) Developing digital worksheets that are more structured and responsive to individual student needs, including the use of dyslexia-friendly fonts (e.g., Open Dyslexic). 4) Interactive, 1–2-minute AI video tutorial sessions to introduce each task.

Data from teacher reflections, student learning outcomes, and AI platform usage logs indicate increased student participation, interest, and accuracy in completing literacy tasks. Teachers also indicated that the design changes helped them provide differentiated interventions tailored to the needs of each student in inclusive classrooms.

IV. DISCUSSION

The findings of this study indicate that the integration of AI into inclusive learning in elementary schools has significant potential to improve the literacy of students with learning disabilities such as dyslexia and dysgraphia. Teachers reported increased student engagement when using AI-based technologies, such as automatic text readers, reading speed adjusters, and automatic handwriting correctors. This provides a more individualized and adaptive learning experience, which has been difficult to provide in resource-constrained inclusive classrooms.

Results from two implementation cycles indicate that instructional design that considers teacher and student feedback can enhance the effectiveness of interventions. The first cycle showed that while AI engaged students, the complexity of using the technology made it difficult for some students. Adjustments in the second cycle, such as the use of supporting visual features, interactive tutorials, and reduced workload, successfully improved the functionality of AI in the context of inclusive classrooms. Thus, AI is not merely a technological tool, but also a medium for translating Universal Design for Learning (UDL) principles into practical practice.

Furthermore, these findings imply that teacher training in the pedagogical use of AI is crucial. Many teachers expressed a need for capacity building in designing and adapting AI technology to the needs of students with special needs. Therefore, teacher professional development must encompass an understanding of both technology and inclusive pedagogy. With proper integration, AI can be a catalyst for more equitable, equal, and equitable education for all students, including those with literacy challenges.

In the Indonesian context, the use of AI in inclusive learning is increasingly relevant, given geographic challenges, student diversity, and the limited number of special educators in remote areas. AI technology can be an alternative solution to bridge the gap in access to inclusive education, especially in elementary schools that lack dedicated assistant teachers. AI-based applications that operate offline or in a hybrid format are also well-suited to the uneven internet infrastructure in many parts of Indonesia.

Research findings indicate that with simple interfaces, localized features such as the use of Indonesian, and ease of content adaptation, AI technology can be functionally integrated into the classroom. Teachers in several public and private elementary schools are able to utilize AI-based applications after receiving brief training, although there is still a need for improved digital literacy. The use of a special dyslexia font, Indonesian-language text narrator, and voice recognition with local accents provide a more contextual learning experience for students with special needs.

However, the use of AI in inclusive education in Indonesia needs to be accompanied by supporting policies from local and central governments, including the provision of hardware and ongoing training for teachers. Furthermore, it is crucial to involve parents and the school community so that the use of technology is not merely instructional but also becomes part of a learning ecosystem that supports the social and emotional development of students with special needs. This indicates that the relevance of AI lies not only in its sophistication, but in its ability to adapt to local needs, values, and conditions.

A key strength of this study is its design-based approach with two implementation cycles, which allowed for contextual adjustments to the instructional design based on teacher reflections and student learning outcomes. The mixed methods study used combining quantitative data from teacher assessments and observations, with qualitative data from interviews and documentation provides a holistic picture of the effectiveness and challenges of using AI in inclusive learning. The success of the second cycle in increasing student engagement is an indicator of the success of a participatory approach that actively involves teachers in the innovation process.

However, this study also has several limitations. First, the number of respondents was limited to 15 teachers from several elementary schools, so generalizing the findings to all of Indonesia requires caution. Second, the study's duration, which only covered two learning cycles, makes the results more exploratory than longitudinal. Third, although several AI-based applications were used, the technology used was relatively simple and did not fully represent the full potential of AI, such as machine learning or automated adaptive learning.

Another limitation is the lack of quantitative, long-term effectiveness testing based on student literacy skills. Although there was an increase in teachers' perceptions of the effectiveness of AI and inclusive classroom management, the measurement of the direct impact on student learning outcomes was still descriptive and requires further study. Therefore, the results of this study are more appropriately viewed as a basis for developing more tested AI-based learning models in the context of inclusive education in Indonesia.

Several international studies have shown results consistent with this research. In Finland, for example, AI is being used to detect reading error patterns in dyslexic students, which are then used to develop data-driven interventions. Similarly, in South Korea, AI has been utilized in language teaching for children with special needs through a speech recognition application capable of providing personalized feedback. As in this study, teacher engagement and adaptive instructional design are key to the success of technology-based interventions.

However, striking differences are evident in the infrastructure and readiness of education systems in these countries compared to Indonesia. In developed countries, AI implementation is supported by advanced hardware, adequate teacher training, and structured educational policy support. In Indonesia, challenges such as limited devices, unequal teacher digital literacy, and disparities in education quality across regions are obstacles that need to be overcome before AI can play an optimal role as in these countries.

However, the unique Indonesian context also opens up opportunities for locally-based AI development. This study's

findings demonstrate that adapting AI features to local language, culture, and student characteristics can yield positive results, even with modest technology. This demonstrates that the effectiveness of AI in inclusive learning is determined not only by technological sophistication, but also by design capabilities that are responsive to the social and educational contexts in which the technology is implemented.

V. CONCLUSION

This research makes a significant contribution to the context of inclusive education in Indonesia by developing and implementing an AI-based learning design tailored for students with dyslexia and dysgraphia. Through a two-cycle Design-Based Research (DBR) approach, it was found that the use of AI improved students' understanding of the material, learning engagement, and writing skills in an inclusive classroom context. Teachers also responded positively to the use of this technology, increasing their capacity to manage AI-based learning. Another contribution is the development of an adaptive instructional design model that can be replicated and further developed by other teachers and educational developers.

Recommendations for Schools, Policymakers, and Educational Technology Developers

Schools are advised to provide intensive and ongoing training to teachers in utilizing AI-based technologies, particularly those designed for specific needs such as dyslexia and dysgraphia. Policymakers need to expand inclusive policies that not only accommodate the needs of students with learning disabilities but also strengthen inclusive and adaptive educational technology infrastructure. Educational technology developers are encouraged to create AI solutions that are not only based on sophisticated algorithms but also consider pedagogical aspects, cultural context, and ease of integration in inclusive elementary classrooms in Indonesia.

Future research should expand the scope of participation to include students from diverse regional backgrounds and with varying levels of special needs to enhance the generalizability of the results. Furthermore, longitudinal studies evaluating the long-term impact of AI integration in inclusive learning are needed. Future directions could also focus on developing technology-based measurement tools to assess the literacy progress of students with special needs in real-time and adaptively. Furthermore, cross-disciplinary collaboration between educators, technologists, and educational psychologists is essential to create an effective and sustainable AI-based learning ecosystem.

ACKNOWLEDGEMENTS

The author expresses his deepest gratitude to the Directorate of Research and Community Service (DPPM) of the Directorate General of Higher Education, Research, and Technology, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, for providing support and funding for this research. This support has enabled the author to develop and implement AI-based learning innovations in the context of inclusive education in elementary schools. Hopefully, the results of this research can make a tangible contribution to the development of inclusive education practices and learning technology in Indonesia.

REFERENCES

- [1] Agheraï, O., Aoula, E. S., & Ahriz, S. (2025). AI Applications in Education: A review. *International Journal of Education and Information Technologies*, 19(March), 33–38. <https://doi.org/10.46300/9109.2025.19.4>
- [2] Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., M. Ameen, B. M., & Khedhir, R. M. (2025). Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health*, 6(March), 100198. <https://doi.org/10.1016/j.glmedi.2025.100198>
- [3] Ahmed, S., Rahman, M. S., Kaiser, M. S., & Hosen, A. S. M. S. (2025). Advancing Personalized and Inclusive Education for Students with Disability Through Artificial Intelligence: Perspectives, Challenges, and Opportunities. *Digital*, 5(2), 1–30. <https://doi.org/10.3390/digital5020011>
- [4] Alkhater, N., Alabbas, A., Zainaldeen, Z., Aldhamin, M., Alwarsh, M., Shubbar, Z., & Zaidan, A. (2025). The Impact of Artificial Intelligence on Students' Learning Experience. *Studies in Systems, Decision and Control*, 568, 75–84. https://doi.org/10.1007/978-3-031-71526-6_7
- [5] Alkhurayyif, Y., & Sait, A. R. W. (2024). A Review of Artificial Intelligence-Based Dyslexia Detection Techniques. *Diagnostics*, 14(21). <https://doi.org/10.3390/diagnostics14212362>
- [6] Bhola, N. (2022). Effect of Text-to-speech Software on Academic Achievement of Students with Dyslexia. *Integrated Journal for Research in Arts and Humanities*, 2(4), 51–55. <https://doi.org/10.55544/ijrah.2.4.45>
- [7] Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and dysgraphia: definition, diagnosis, and management. *Translational Pediatrics*, 9(3), S46–S54. <https://doi.org/10.21037/TP.2019.11.01>
- [8] Clemens, N. H., & Vaughn, S. (2023). Understandings and Misunderstandings About Dyslexia: Introduction to the Special Issue. *Reading Research Quarterly*, 58(2), 181–187. <https://doi.org/10.1002/rrq.499>
- [9] Creswell, J. W. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage.
- [10] Daley, S. G., & Rappolt-Schlichtmann, G. (2018). Stigma Consciousness Among Adolescents With Learning Disabilities: Considering Individual Experiences of Being Stereotyped. *Learning Disability Quarterly*, 41(4), 200–212. <https://doi.org/10.1177/0731948718785565>
- [11] Gheysens, E., Griffl-Freixenet, J., & Struyven, K. (2023). Differentiated Instruction as an Approach to Establish Effective Teaching in Inclusive Classrooms. In *Effective Teaching Around the World: Theoretical, Empirical, Methodological and*

- Practical Insights* (pp. 677–689). https://doi.org/10.1007/978-3-031-31678-4_30
- [12] Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences*, 13(12). <https://doi.org/10.3390/educsci13121216>
- [13] Iliska, D., & Gudoniene, D. (2025). Sustainable Technology-Enhanced Learning for Learners with Dyslexia. *Sustainability* (Switzerland), 17(10). <https://doi.org/10.3390/su17104513>
- [14] Khatun, M., Islam, R., Kumar, S., Hossain, R., & Mani, L. (2024). The Impact of Artificial Intelligence on Educational Transformation: Trends and Future Directions. *Journal of Information Systems and Informatics*, 6(4), 2347–2373. <https://doi.org/10.51519/journalisi.v6i4.879>
- [15] Kim, S., Cho, S., Kim, J. Y., & Kim, D. J. (2023). Statistical Assessment on Student Engagement in Asynchronous Online Learning Using the k-Means Clustering Algorithm. *Sustainability* (Switzerland), 15(3). <https://doi.org/10.3390/su15032049>
- [16] Kunwar, R., & Sapkota, H. P. (2022). An Overview of Dyslexia: Some Key Issues and Its Effects on Learning Mathematics. *Turkish International Journal of Special Education and Guidance & Counselling* (TIJSEG), 11(2), 82–98.
- [17] Martan, V., Mihić, S. S., & Čepić, R. (2023). Teachers' Knowledge about Students with Dyslexia and Professional Development. *European Journal of Contemporary Education*, 12(2), 535–552. <https://doi.org/10.13187/ejced.2023.2.535>
- [18] Mounkoro, I., Ubras, A. D., Cadelina, F. A., Saleem, M. U., Maryam, S., Almagharbeh, W. T., Abdul Majeed, A., & Hidayatallah, Z. (2025). Integrating AI in Classroom Management: Improving Educational Efficiency and Teacher Workflows. *Journal of Information Systems Engineering and Management*, 10(45s), 823–834. <https://doi.org/10.52783/jisem.v10i45s.9033>
- [19] Nurullayevna, S. I., Shahina Ismatjon, N., & Muxlisa Musulmonqul, P. (2025). Inclusive Education for Children with Disabilities: Overcoming Barriers and Unlocking Opportunities. *Indonesian Journal of Community and Special Needs Education*, 6(1), 1–8.
- [20] Onesi-Ozigagun, O., Yinka James Ololade, Nsiong Louis Eyo-Udo, & Damilola Oluwaseun Ogundipe. (2024). Revolutionizing Education Through Ai: A Comprehensive Review of Enhancing Learning Experiences. *International Journal of Applied Research in Social Sciences*, 6(4), 589–607. <https://doi.org/10.51594/ijarss.v6i4.1011>
- [21] Pagliara, S. M., Bonavolontà, G., Pia, M., Falchi, S., Zurru, A. L., Fenu, G., & Mura, A. (2024). The Integration of Artificial Intelligence in Inclusive Education: A Scoping Review. *Information* (Switzerland), 15(12). <https://doi.org/10.3390/info15120774>
- [22] Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? A review of US research. *Journal of Professional Capital and Community*, 4(4), 286–308. <https://doi.org/10.1108/JPCCC-12-2018-0032>
- [23] Pragasthi, M., & Vaishnavi, Ms. N. (2025). AI – Powered Learning Disability Detection and Classification System. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 11(2), 1191–1196. <https://doi.org/10.32628/cseit5112454>
- [24] Priyadharsini, V., & Sahaya Mary, R. (2024). Universal Design for Learning (UDL) in Inclusive Education: Accelerating Learning for All. *Shanlax International Journal of Arts, Science and Humanities*, 11(4), 145–150. <https://doi.org/10.34293/sijash.v11i4.7489>
- [25] Rahim, N., Mokmin, N. A. M., & Wang, J. (2025). Empowering students with dysgraphia: the educational benefits of assistive technologies. *Disability and Rehabilitation: Assistive Technology*, 20(7), 2025–2036. <https://doi.org/10.1080/17483107.2025.2493737>
- [26] Ranbir, D. (2024). Inclusive Education Practices for Students with Diverse Needs. *Innovative Research Thoughts*, 10(1), 142–146. <https://doi.org/10.36676/irt.v10.i1.1405>
- [27] Rokade, D., Jabde, M. K., & Patil, C. H. (2024). Screening Application of Dyslexia and Dysgraphia Using Cognitive AI. *2024 10th International Conference on Smart Computing and Communication (ICSCC 2024)*, 182–186. <https://doi.org/10.1109/ICSCC62041.2024.10690576>
- [28] Scott, E. E., Wenderoth, M. P., & Doherty, J. H. (2020). Design-based research: A methodology to extend and enrich biology education research. *CBE Life Sciences Education*, 19(3), 1–12. <https://doi.org/10.1187/cbe.19-11-0245>
- [29] Sharma, S., Tomar, V., Yadav, N., & Aggarwal, M. (2023). Impact of AI-based special education on educators and students. In *AI-Assisted Special Education for Students With Exceptional Needs* (pp. 47–66). IGI Global. <https://doi.org/10.4018/979-8-3693-0378-8.ch003>
- [30] Snowling, M. J., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: past, present and future. *Oxford Review of Education*, 46(4), 501–513. <https://doi.org/10.1080/03054985.2020.1765756>
- [31] Tweeten, J., & Hung, W. (2023). Design-based Research Method in PBL / PjBL: A Case in Nursing Education. *The Interdisciplinary Journal of Problem-Based Learning*, 17(2).
- [32] Yusyac, R., Muslem, A., & Yasin, B. (2021). Using Contextual Teaching and Learning (CTL) Approach to Improve Students' Speaking Ability. *English Education Journal*, 12(3), 460–476.
- [33] Zaini, N., Jamil Azhar, S. B. H., Atan, M. A., Norwahi, N. A., Abdul Hamid, A. A., & Mazlan, N. A. (2024). Social Stigma Against Dyslexics: Exploring Public's Knowledge and Belief. *International Journal of Academic Research in Business and Social Sciences*, 14(6), 817–826. <https://doi.org/10.6007/ijarbss.v14-i6/21706>
- [34] Zhu, Y. (2024). The Impact of AI-Assisted Teaching on Students' Learning and Psychology. *Journal of Education, Humanities and Social Sciences*, 38, 111–116. <https://doi.org/10.54097/k7a37d11>



Sri Sukasih, one of the expert lecturers in elementary education at Universitas Negeri Semarang (Semarang State University) Central Java, Indonesia. She is an active lecturer and scientific writer who has published 69 papers in national and international journals. The 2025 scientific paper that has been published is entitled "Improving Mindful, Meaningful, and Joyful Learning Design Skills Through the Utilization of Virtual Laboratories to Realize Deep Learning at the KKG Gugus Kemuning Elementary School, Semarang City". She is currently still an Associate Professor at Universitas Negeri Semarang (UNNES). She can be contacted via email: srisukasih@mail.unnes.ac.id



Atip Nurharini, one of the lecturers in elementary education at Universitas Negeri Semarang (Semarang State University) Central Java, Indonesia. She is a graduate of PGSD (Department of Elementary Education). Her scientific wealth consists of 74 publications in national journals and international scientific journals. She can be contacted at email: atip.nurharini@mail.unnes.ac.id



Andarini Permata Cahyaningtyas is an Elementary School Teacher Education Lecturer with a master's degree from Universitas Negeri Yogyakarta and a bachelor's degree from Universitas Negeri Yogyakarta. Since 2022, she has joined the Elementary Teacher Education Study Program at Universitas Negeri Semarang, Indonesia. Her research interests include children's literature, language for children, and elementary education. She can be contacted via email: andarinipermata@mail.unnes.ac.id

The Role of Tone in Dubbing Children's Cartoons: A Case Study of *Timon & Pumbaa* Series

Alshaymaa Y. Alharbi
Applied College, Taibah University, Madinah, Saudi Arabia

Abstract—Translation plays a vital role in crossing cultural boundaries, but effectively conveying the tone and spirit of original works is challenging. This study analyzes how the vocal tone of different voice actors impacted audience perception of the animated character Timon from Disney's *The Lion King* franchise when dubbed into Arabic. Specifically, it examines audience feedback on the portrayals by Mohamed Henedy and Nasser Shaheen in the *Timon and Pumbaa* television series. A questionnaire was designed to identify 53 faculty members' perceptions on the voice actors' ability to capture Timon's comedic wit and sarcastic personality through vocal tone. Results show Henedy's energetic and humorous delivery strongly aligns with Timon's characterization, making his performance more engaging and memorable. Though Shaheen succeeded in conveying sarcasm and wit, his portrayal was perceived as less well-rounded. The results emphasize vocal tone's pivotal role in authentically transporting an animated character's essence cross-culturally. Matching the spirit and nuances of the original voice acting helps elevate competent dubbing to truly iconic status. The study provides unique insights into how attention to tone impacts audience reception and popularity of translated animated works.

Index Terms—translation, voice acting, vocal tone, characterization, dubbing

I. INTRODUCTION

Translation is indispensable in our everyday life for its being the most diverse means of human communication. In this age of globalization and widespread immigration, the need for translation is increasing due to the continuous expansion of trade, science, culture, technology and so on. Translation is a field of contrastive linguistics since it is associated with at least two languages and their cultures. It is the process of conveying a message by transferring it from one system of language into another. That is to say, it is the technique used to transfer and/or transform the meaning of a source language into a target language (Akan et al., 2019). However, although the translator's voice is everywhere, the translator is nowhere to be seen. This is the paradox of translation: in presenting his translation, the translator is a speaker, but in this very same situation, he is not the speaker. His voice is not really his voice. It is the voice of someone else, as it has no original status (Suchet, 2013).

Performing a role that has been translated into another language, however, brings a different aspect to the meaning an actor may convey in the dubbed version. Timon and Pumbaa are an animated meerkat and warthog duo introduced in Disney's 1994 animated film *The Lion King* and its franchise (Lesnick, 2017). A special animated series was made for them called *Timon and Pumbaa*, which was dubbed into Arabic, and the role of Timon was played by the following actors: Muhammad Henedy and Nasser Shaheen, respectively. Tone means the feeling and meaning of a melody, that essential core so hard to put into words and most viewers are affected by the tone of the translation of a particular person. There are those who like the tone of a specific actor and do not want to accept another actor in the role (Mazanec, 2011). This research will study the effect of tone on the performance of the Timon character by Mohamed Henedy and then Nasser Shaheen in the *Timon & Pumbaa* TV series. The main research question for this study is as follows:

What is the effect of tone on the translation of the Timon character by Mohamed Henedy and Nasser Shaheen in the *Timon & Pumbaa* TV series?

In order to have an optimal result, the research scope is limited to investigating the different effects of tone in the translation of the Timon character between the two actors. This research can thus provide more understanding of the effect and importance of tone in presenting dubbed versions of children's programmes.

II. LITERATURE REVIEW

Translation of animated works aimed at a young audience requires adept adaptation to convey tone appropriately across cultures. Dubbing and vocal performance play a particularly crucial role in shaping children's impressionable connections with on-screen characters (Chaume, 2012). The complex art of dubbing children's media has received increasing scholarly attention since studios recognize the importance of tailoring content for global audience while preserving narrative intent and impact (O'Connell, 2003). Dubbing offers the capacity to reinterpret dialogue, timing and vocal tone while matching mouth movements more closely. Meanwhile, subtitling allows preservation of original voice performances but requires

greater editing to condense and simplify language (Perego, 2008). The chosen translation approach shapes fundamental aspects of the viewing experience, from humor to characterizations. As Matamala (2009) notes, “Translating for children represents first of all a great responsibility. Their cognitive skills are still developing, so we have to be very careful with issues like vocabulary, syntax, semantics, even intonation” (p. 217).

Dubbing and revoicing provide opportunities to recalibrate animated content, adjusting cultural references, wordplay, and idioms that may not resonate across languages (Bucaria & Scifo, 2009). The meaning behind a line of dialogue takes priority over verbatim translation. As Chaume (2012) observes, “The main aim of revoicing is to produce in the target language the same effect the original product had on the original addressees” (p. 22). Translating for children incorporates additional ethical considerations regarding what impressions and values we seek to instill in young global citizens through shared media. When translating animated films and series for children, the lexical choices applied to central characters are impactful in shaping audience perceptions (Bruti, 2009). Characters’ distinctive ways of speaking inform their identities. As O’Hagan (2009) explains, “The translation of dialect and sociolects is an important consideration in rendering animated films, where speech defines character to a great extent” (p. 207). The translator must determine which culturally-specific attributes to retain versus adapting and reconstructing animated personalities for the target culture (Chaume, 2021). Regional accents and quirks from the source material may need to be replaced by more relatable signifiers.

Perez-Gonzalez (2014) argues that domestication often predominates in the mainstream dubbing industry, orienting translated works toward the dominant cultural perspectives of the target audience. Accents and word choices deemed too foreign or niche typically undergo generalization. This prioritization of familiarity and supposed universality risks stripping away meaningful diversity in modes of expression. Perez-Gonzalez advocates audiovisual translation approaches that embrace linguistic hybridity, including creative use of accents to signal characters’ cultural backgrounds in a manner accessible to children (Perez-Gonzalez, 2014). Effective dubbing requires vocal performances that capture the essence of animated characters. As Barbañán (2019) notes, “The voice characterizes our personas, and children need to recognize their fictional heroes and heroines” (p. 154). Vocal pitch, diction, rhythm, and affect all contribute to the audience’s impressions. Casting voice actors who can relate to young viewers is crucial when reinventing roles.

Perego (2008) indicates the importance of humor in achieving functional equivalence between the source text and the dubbed version. Puns, wordplay, and jokes rooted in linguistic and cultural context often necessitate ingenious reworking. As Bruti (2009) states, “The translation of wordplay, which abounds in children’s films, is an extremely challenging task, not only because of its strong dependence on the source text, but also because it is strictly connected to the oral dimension” (p. 26). Animated comedies rely heavily on vocal delivery and timing. The complexity of translating humor while maintaining spontaneity underscores the creative problem-solving that dubbing requires.

Rowe (2018) highlights Disney’s global franchise model and the company’s historic prioritization of domestication in translation. For decades, Disney typically adhered to strict dubbing practices eschewing foreign accents in their animated films. Rowe traces a gradual evolution toward incorporating more diverse and hybridized voices, as in the case of the Spanglish-voiced character Mirabel in *Encanto* (2021). Rowe argues that growing acknowledgment of multicultural audience encourages inclusion of “different global, national and local norms” in dubbing major productions (Rowe, 2018, p. 156). Mainstream studios now demonstrate greater interest in adapted translations that thoughtfully balance localization with authentic representation.

The translation of children’s media blends craft, care, and responsibility in reinterpreting resonant stories for new contexts. Dubbing and subtitling involve creative adaptation as well as adjustments driven by industry priorities and target audience. As O’Connell (2003) summarizes: “Good translation for children requires something that everyone concerned with their development constantly strives to encourage: productive, intelligent interaction with the multi-faceted world around them” (p. 215). The significance of mass media in shaping children’s perspectives imbues translation of youth-oriented content with added meaning. Analytical scholarship has continued elucidating the layered variables and interests that come into play when reinventing animated worlds through translated dialogue.

A. Foundational Work on Animated Translation

Early foundational work by Chaume (2012) thoroughly examines the multifaceted challenges of translating audiovisual content for children. He outlines the strategies used to bridge cultural gaps and adapt references to suit young viewers’ knowledge. Chaume (2012) notes that translators must creatively rewrite jokes, wordplay, songs, and dialogue to entertain and educate children while respecting censorship norms. Vocal tone and performance are critical for bringing animated characters to life consistently across languages. He advocates preserving the original spirit and conversational flow rather than literal accuracy. Building on Chaume’s (2012) research, Igareda (2016) identifies domesticating techniques used when translating Disney musical films into Spanish. Expanded dialogue and localized song lyrics were utilized to facilitate natural timing with characters’ mouth movements and match personalities. Igareda (2016) argues this stylistic freedom is essential for conveying the right comedic or dramatic tone in translated musical numbers. Her analysis showcases the priority of maintaining characters’ essence over source text fidelity in functional children’s adaptation.

Perego (2008) narrows the lens to examine translation strategies for wordplay and humor in Disney films. He unpacks the multifaceted decisions involved in balancing enjoyable amusement for young audience with faithfulness to the original text. Perego (2008) notes timing, word choice, and vocal inflection as paramount in transmitting a character’s humorous

tone. He concludes the translator's creativity, comic sensibility and inner child mindset are vital for amusing dubbed versions that capture personality.

The dubbing of children's literature plays a crucial role in the educational process. A study by Saed et al. (2024) analyzes how dubbing animated films from English into Arabic for educational purposes stimulates children, focusing on both educational and entertainment aspects. Dubbing is more than simply translating words; it involves simplifying the content by adjusting the tone, humor, and teaching style. This process seeks to balance educational objectives with linguistic and cultural requirements. This study highlights the practical challenges of preserving the expressiveness and emotion that the original text intends to convey. One of these challenges is that translators must balance accuracy with adapting the content to children's cognitive level to ensure effective learning.

Mohammedi (2023) focuses on maintaining the original intonation of the source language when watching foreign-language animated films with Algerian children. The study concludes that the appropriate intonation contributes to the authenticity of the work. Inconsistent intonation due to cultural or rhythmic differences can hinder children's comprehension. This study confirms that children's engagement and comprehension are closely linked to the tonal nuances of the dubbed voice. This highlights that dubbing is more than simply transcribing words. However, it is the recreation of all the phonetic characteristics necessary to create relevant and effective content for the target audience.

B. Reception Studies and Audiences' Perspectives

Scholars have also increasingly considered audience perspectives, studying how translation techniques shape children's reception and engagement with animated stories. González Vera (2018) investigates how Spanish dubbing versus subtitled versions impacted perceptions of animated films. Surveys revealed most participants preferred dubbing, citing superior enjoyment, comfort, and connection with characters. The quality of voice acting performance emerged as a key factor driving immersive appeal for children. This underscores vocal tone's influence in bringing animated worlds to life across cultures. In a broader study on adults' attitudes, Igareda (2011) similarly finds out a strong inclination toward dubbing among Spanish audiences based on benefits like humor preservation. Though some concerns emerged around loss of original voices, most felt high-quality dubbing enhanced comprehension and entertainment. Igareda (2011) argues exposure and habit play a key role in normalized dubbing preferences in Spain's audiovisual landscape. However, research on audiences outside dubbing cultures shows contrasting preferences. Wissmath and Weibel's (2012) research in Switzerland reveals most viewers preferred subtitles over dubbed versions due to original voices' authenticity. They also note an association of dubbing with children's content, suggesting target age shapes audience expectations. Although inconclusive, these studies showcase the nuanced impacts of translation approaches on engagement.

C. Debates on Domestication and Foreignization

Some scholars have critiqued excessive domesticating techniques in translating children's media as potentially stripping away cultural elements. Marine-Roig (2015) surfaces tensions between domestication for fluency and foreignization for authenticity regarding Disney comic translations into Catalan. She argues domestically oriented approaches can erase valuable cultural context clues for young readers. Marine-Roig (2015) calls for balanced translation that stimulates children's intercultural learning while remaining engaging. Similarly emphasizing source fidelity, Martínez Sierra (2004) advocates measured functional translation that adapts texts when necessary for comprehension and enjoyment while avoiding extraneous alterations that erase cultural markers. Martínez Sierra asserts, "translators should not assume children are incapable of dealing with new concepts" (p. 177). She argues respecting source materials facilitates exposure to enriching new ideas. O'Connell (2003) extends these concerns around excessive domestication to audiovisual translation, suggesting it can result in homogeneous international versions devoid of culturally educational elements. However, she acknowledges entertainment value as the prime aim for children's media translation, a nuanced balancing act.

Wang et al. (2024) conducted a comprehensive review of research on dubbing as a foreign language learning tool between 2017 and 2023. By analyzing students' attitudes and speaking skills, they conclude that the tone and emotional quality of dubbing increase motivation and engagement in learning, as well as improving the authenticity of the text. Furthermore, students rate dubbing positively, indicating that the natural intonation can help them improve their pronunciation and confidence. However, dubbing can present certain challenges, such as poor sound quality and clarity, which can be overcome by focusing on appropriate intonation.

D. Multimodal Considerations Beyond Language

Beyond verbal language, scholars emphasize multimodal factors that shape tone and characterization in translating children's animation. As Mera (1999) notes, music narratively embodies emotion and identity in successfully conveying tone and personality requires skillfully translating lyrical meaning while maintaining musicality. Songs act as vehicles for mood and character expression. Maier (2019) similarly discusses the interconnectedness of visuals, dialogue, and sound in animations' holistic storytelling. Therefore, she suggests translators must attend to all channels together when adapting tone, humor, and pacing for children. Maier provides guidelines for multimedia translation fidelity centered on consistent characterization. Paratextual elements also influence reception. Nicolas (2020) explores Spanish title translations for Disney-Pixar films aiming at piquing children's interest. Strategies like alliteration, rhyme, and semantic

transfer were used to evoke similar reactions as original titles. This demonstrates the subtle ways translators shape engagement surrounding children's media.

Wit and attention are two arts, and mastering them can generate significant benefits and have a major impact in the educational field. Alsuhaime (2024) analyzes the influence of wit and attention on dialogue dubbing in children's films, using the television series "The Loud House" as a case study. This study, based on Brown and Levinson's theory, found that the more accurate the dubbing pitch, the better young children understood and interacted with social relationships, interactions, and etiquette in the new language.

E. Accents and Representation

Lastly, translation choices around accents demand consideration regarding potential stereotyping or underrepresentation. As Wikholm (2012) observes, foreign accents in animation dubbing frequently index comic antagonists. He highlights concerning implications around using accents to signal negative traits and warns against undermining nuanced characterization. Guidelines thus emphasize using accents intentionally, not as shorthand for foreign villain coding. Overall, research highlights the multifaceted and nuanced art of translating children's animated media to entertain and enrich young minds across cultures.

Saiddina and Darma's (2024) qualitative study takes a different approach when studying the impact of animation on English language development in bilingual children. Their results show that variations and repetitions in dubbing pitch promote contextual understanding and phonological awareness. This study recommends that parents watch animation with their children to improve language acquisition. Notably, this study examines not only the dubbing itself but also variations in pitch and rhythm, which can help children understand the linguistic patterns of the second language environment.

From language adaptation to song lyrics to dialogue delivery, maintaining engaging tone and characterization emerges as crucial yet complex. Continued empirical reception studies can provide unique insights into how translational approaches and vocal performance elements successfully captivate and connect with diverse young audiences. This research aims to contribute much-needed focus on Arabic translation and dubbing practices in children's animation, which is: an understudied context.

III. METHODOLOGY

This research is based on a quantitative method, using a questionnaire as a tool. The study sample includes viewers who watched the animated series *Timon and Pumbaa* with the dubbing of Timon's character by Mohamed Hendy and Nasser Shaheen to clarify their opinions about the differences in tone between the two voice actors. The questionnaire was designed by the researcher using a five-point Likert scale. Data were collected from this questionnaire using an SPSS program v.25 for statistical analysis.

The research sample comprised 53 faculty members specializing in linguistics at Saudi universities. The data was collected by categorizing the participants into three age groups: 26-35, 36-45, and older than 45. The number of male and female faculty members was recorded for each age group. Table 1 shows the gender distribution and age groups, and Figures 1 and 2 show the percentage of gender distribution and age groups respectively.

TABLE 1
GENDER DISTRIBUTION AND AGE GROUPS

Age	Male	Female	Total
26-35	2	5	7
36-45	15	9	24
Older than 45	16	6	22
Total	33	20	53

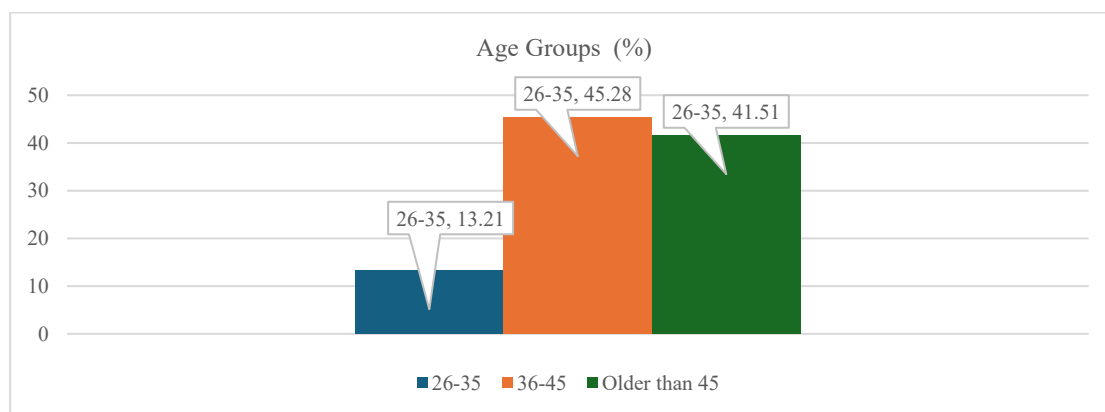


Figure 1. Percentage of Age Groups

Table 1 and Figure 1 reveal interesting patterns in terms of gender distribution and age groups. Among the three age categories, the highest number of faculty members falls within the 36-45 age range with 24 individuals (45.28%). This group is closely followed by those older than 45 age group, comprising 22 faculty members (41.51%). The youngest age group, 26-35, has the fewest members, with a total of 7 (13.21%). Figure 2 shows the percentage of gender distribution.

From Figure 2 and Table 1, it can be concluded that male faculty members represent the majority across all age groups, totaling 33 individuals (62%), while female faculty ones account for a smaller proportion, with 20 individuals (38%). These findings highlight the need for further exploration to understand the factors that may contribute to the underrepresentation of female linguistics faculty members, particularly in younger age groups.

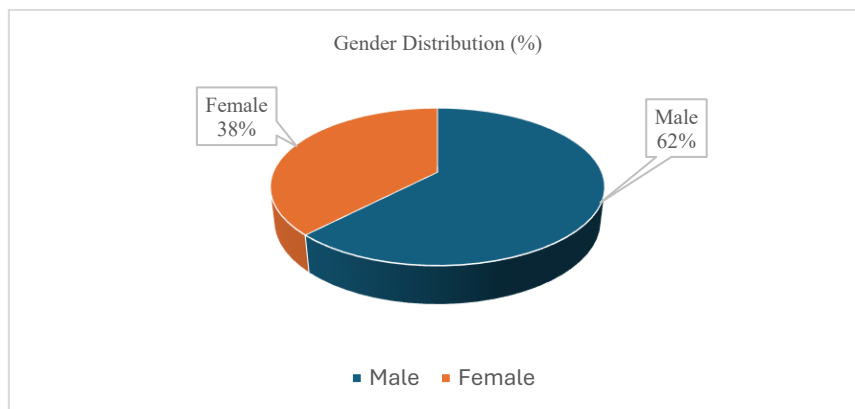


Figure 2. Percentage of Gender Distribution

In order to assess the reliability of the research tool (questionnaire), Cronbach's Alpha (α) formula was employed on the sample. The data presented in Table 2 clearly shows that the study variables have a high level of overall stability, with a stability coefficient of 0.910 for all 20 items in the questionnaire. The stability of the individual variables ranges from a minimum of 0.890 to a maximum of 0.911. This demonstrates that the questionnaire is highly stable and can be confidently used in the practical implementation of the study. Table 2 shows the stability scores of the research tool.

TABLE 2
CRONBACH'S ALPHA COEFFICIENT FOR ASSESSING THE CONSISTENCY OF THE RESEARCH TOOL

Dimension	Number of Items	Topic Stability
Tone of Mohamed Henedy's voice acting	10	0.890
Tone of Nasser Shaheen's voice acting	10	0.911
Overall stability of the questionnaire	20	0.910

IV. RESULTS AND DISCUSSION

The descriptive analysis provides an overview of the responses for each question item in the questionnaire. It includes the mean and standard deviation values, which summarize the central tendency and dispersion of the responses. Table 3 shows the descriptive statistics for the first dimension, which covers the tone of Mohamed Henedy's voice acting. It includes the mean and standard deviation for each question item related to Henedy's vocal performance as Timon.

TABLE 3
DESCRIPTIVE ANALYSES OF THE FIRST DIMENSION (TONE OF MOHAMED HENEDY'S VOICE ACTING)

#	Dimension Items	Mean	Std. Deviation
1	Mohamed Henedy's voice acting tone was energetic and enthusiastic.	3.77	0.776
2	Mohamed Henedy's voice tone conveyed the comedic timing well.	4.00	0.832
3	Mohamed Henedy's tone captured the sarcastic nature of Timon's lines.	3.87	0.708
4	Mohamed Henedy's tone was engaging and expressive for the character.	3.87	0.652
5	Mohamed Henedy's tone matched Timon's personality based on the animation.	3.92	0.851
6	Mohamed Henedy's comedic tone fits the jokes and humor of the show.	3.81	0.681
7	Mohamed Henedy's tone enhanced the meaning of Timon's lines.	3.94	0.770
8	Mohamed Henedy's tone made the character more enjoyable to watch.	3.77	0.750
9	Mohamed Henedy's tone worked well for the witty nature of Timon.	3.81	0.652
10	Overall, Mohamed Henedy's tone suited the character of Timon.	3.94	0.795

The descriptive statistics in Table 3 and mean results in Figure 3 show that respondents rated Mohamed Henedy's voice acting tone positively overall for suiting the character of Timon. The mean scores ranged from 3.77 to 4.00 on the 5-point Likert scale, indicating moderate to high agreement. In particular, Henedy's tone was seen as effectively conveying comedic timing, matching Timon's animated personality, enhancing the meaning of dialogue, and capturing the sarcastic nature of the character. The high mean scores and relatively low standard deviations suggest general consensus among respondents regarding the strength and appropriateness of Henedy's vocal performance for the role of Timon.

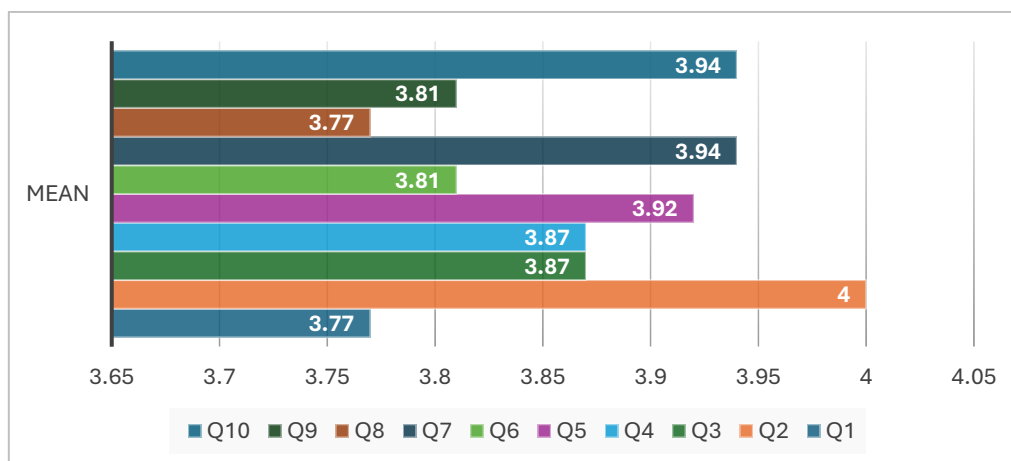


Figure 3. Mean Results of the First Dimension (Tone of Mohamed Henedy's Voice Acting)

Table 4 shows the descriptive statistics for the second dimension, which covers the tone of Nasser Shaheen's voice acting. It includes the mean and standard deviation for each question item related to Shaheen's vocal performance as Timon.

TABLE 4
DESCRIPTIVE ANALYSES OF THE SECOND DIMENSION (TONE OF NASSER SHAHEEN'S VOICE ACTING)

#	Dimension Items	Mean	Std. Deviation
1	Nasser Shaheen's voice acting tone was energetic and enthusiastic.	3.72	.769
2	Nasser Shaheen's voice tone conveyed the comedic timing well.	3.72	.794
3	Nasser Shaheen's tone captured the sarcastic nature of Timon's lines.	3.74	.684
4	Nasser Shaheen's tone was engaging and expressive for the character.	3.57	.721
5	Nasser Shaheen's tone matched Timon's personality based on the animation.	3.83	.700
6	Nasser Shaheen's comedic tone fits the jokes and humor of the show.	3.58	.819
7	Nasser Shaheen's tone enhanced the meaning of Timon's lines.	3.57	.694
8	Nasser Shaheen's tone made the character more enjoyable to watch.	3.58	.719
9	Nasser Shaheen's tone worked well for the witty nature of Timon.	3.68	.872
10	Overall, Nasser Shaheen's tone suited the character of Timon.	3.47	.775

The descriptive statistics in Table 4 and mean results in Figure 4 summarize responses regarding Nasser Shaheen's vocal performance as Timon. The mean scores ranged from 3.47 to 3.83 on the 5-point Likert scale, reflecting slightly lower levels of agreement compared with ratings for Mohamed Henedy. However, Shaheen's tone was still perceived positively overall in terms of suiting Timon's character, with strengths noted in conveying enthusiasm, matching the animated personality, and capturing sarcasm. The relatively higher standard deviations indicate more variability in responses concerning Shaheen compared with Henedy. Overall, the descriptive analysis suggests Shaheen's voice acting was seen as appropriate for Timon by respondents, but to a slightly lesser extent compared with Henedy based on the lower mean scores.

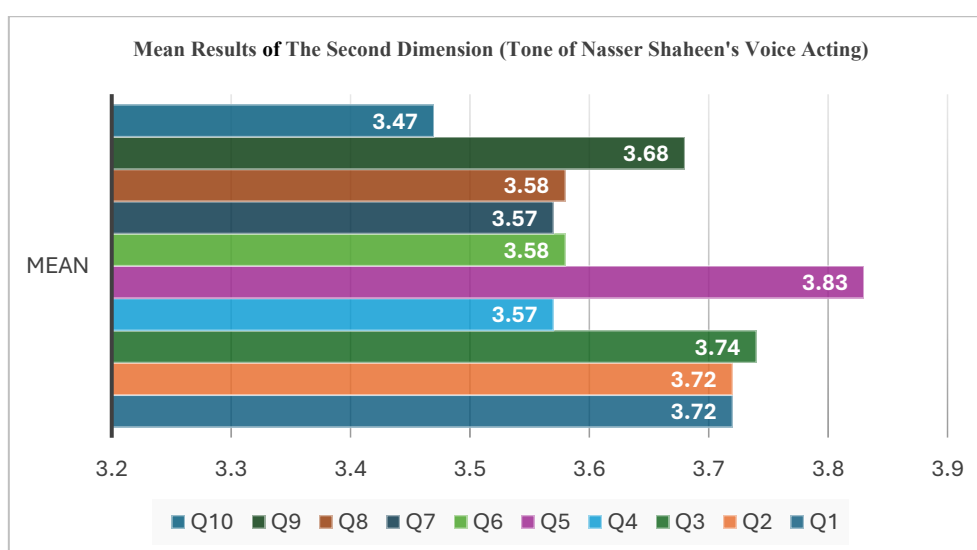


Figure 4. Mean Results of the Second Dimension (Tone of Nasser Shaheen's Voice Acting)

The correlation analysis examines the statistical relationship between each questionnaire item and the overall score for the corresponding dimension. It provides Pearson correlation coefficients to quantify the strength and direction of the associations. The correlation coefficients range from -1 to 1, with values close to 1 indicating a strong positive correlation, values close to -1 indicating a strong negative correlation, and values near 0 suggesting a weak or negligible correlation. The correlation analysis indicates how closely each aspect of voice acting tone aligns with the overall assessment for that voice actor. It reveals which specific elements of the vocal performance are most closely related to the overall tone suiting the character of Timon. In Table 5, correlation coefficients between each item within the first dimension (Tone of Mohamed Henedy's voice acting) and the total score of that dimension can be seen.

TABLE 5
PEARSON CORRELATION COEFFICIENT RESULTS OF THE FIRST DIMENSION (TONE OF MOHAMED HENEDY'S VOICE ACTING)

#	Dimension Items	Correlation Coefficients	Value of Significance
1	Mohamed Henedy's voice acting tone was energetic and enthusiastic.	.754**	.000
2	Mohamed Henedy's voice tone conveyed the comedic timing well.	.744**	.000
3	Mohamed Henedy's tone captured the sarcastic nature of Timon's lines.	.629**	.000
4	Mohamed Henedy's tone was engaging and expressive for the character.	.711**	.000
5	Mohamed Henedy's tone matched Timon's personality based on the animation.	.684**	.000
6	Mohamed Henedy's comedic tone fits the jokes and humor of the show.	.740**	.000
7	Mohamed Henedy's tone enhanced the meaning of Timon's lines.	.706**	.000
8	Mohamed Henedy's tone made the character more enjoyable to watch.	.683**	.000
9	Mohamed Henedy's tone worked well for the witty nature of Timon.	.695**	.000
10	Overall, Mohamed Henedy's tone suited the character of Timon.	.766**	.000

**. Correlation is significant at the 0.01 level (2-tailed)

The correlation analysis in Table 5 demonstrates statistically significant positive correlations between all the questionnaire items concerning Mohamed Henedy's vocal performance and the overall score for his voice acting tone dimension. The strongest correlations are with the items assessing whether Henedy's tone conveyed comedic timing (0.744), matched Timon's personality (0.684), enhanced the meaning of dialogue (0.706), and overall suited the character (0.766). The weakest correlation is with the item regarding capturing sarcasm (0.629). Overall, the consistently high correlation coefficients suggest that Henedy's vocal tone various traits, such as energy (0.754), comedic timing (0.744), meaning (0.706), and enjoyability (0.683), all contribute to the perception that his voice acting effectively fits the character of Timon. No single item stands out as irrelevant.

Table 6 shows the correlation coefficients between each item within the second dimension (Tone of Nasser Shaheen's voice acting) and the total score of that dimension.

TABLE 6
PEARSON CORRELATION COEFFICIENT RESULTS OF THE SECOND DIMENSION (TONE OF NASSER SHAHEEN'S VOICE ACTING)

#	Dimension Items	Correlation Coefficients	Value of Significance
1	Nasser Shaheen's voice acting tone was energetic and enthusiastic.	0.749**	0.000
2	Nasser Shaheen's voice tone conveyed the comedic timing well.	0.709**	0.000
3	Nasser Shaheen's tone captured the sarcastic nature of Timon's lines.	0.696**	0.000
4	Nasser Shaheen's tone was engaging and expressive for the character.	0.707**	0.000
5	Nasser Shaheen's tone matched Timon's personality based on the animation.	0.683**	0.000
6	Nasser Shaheen's comedic tone fits the jokes and humor of the show.	0.808**	0.000
7	Nasser Shaheen's tone enhanced the meaning of Timon's lines.	0.711**	0.000
8	Nasser Shaheen's tone made the character more enjoyable to watch.	0.693**	0.000
9	Nasser Shaheen's tone worked well for the witty nature of Timon.	0.860**	0.000
10	Overall, Nasser Shaheen's tone suited the character of Timon.	0.814**	0.000

**. Correlation is significant at the 0.01 level (2-tailed)

For Nasser Shaheen's vocal performance, Table 6 also shows statistically significant positive correlations between all questionnaire items and the overall tone score. The strongest correlations are with items concerning comedic timing (0.808), wit (0.860), and overall character suitability (0.814). The weakest correlations are for items regarding energy (0.749), meaning (0.711), and enjoyability (0.693).

In general, Shaheen's effectiveness at conveying Timon's wit (0.860) and matching the comedic timing (0.808) of the show appears most integral to his vocal tone suiting the character. Traits like energy (0.749) and meaning enrichment (0.711) correlate less strongly, indicating they contribute less to Shaheen's successful portrayal based on audience perceptions. Table 7 and Figure 5 compare the findings of Table 4 and Table 5 as follows:

TABLE 7
PEARSON CORRELATION COEFFICIENT RESULTS OF THE QUESTIONNAIRE

#	Voice acting tone	Mohamed Henedy	Nasser Shaheen
1	Energetic and enthusiastic.	.754**	.749**
2	Conveyed the comedic timing well.	.744**	.709**
3	Captured the sarcastic nature of Timon's lines.	.629**	.696**
4	Engaging and expressive for the character.	.711**	.707**
5	Matched Timon's personality based on the animation.	.684**	.683**
6	Fit the jokes and humor of the show.	.740**	.808**
7	Enhanced the meaning of Timon's lines.	.706**	.711**
8	Made the character more enjoyable to watch.	.683**	.693**
9	Worked well for the witty nature of Timon.	.695**	.860**
10	Overall, Voice acting tone suited the character of Timon.	.766**	.814**

** . Correlation is significant at the 0.01 level (2-tailed)

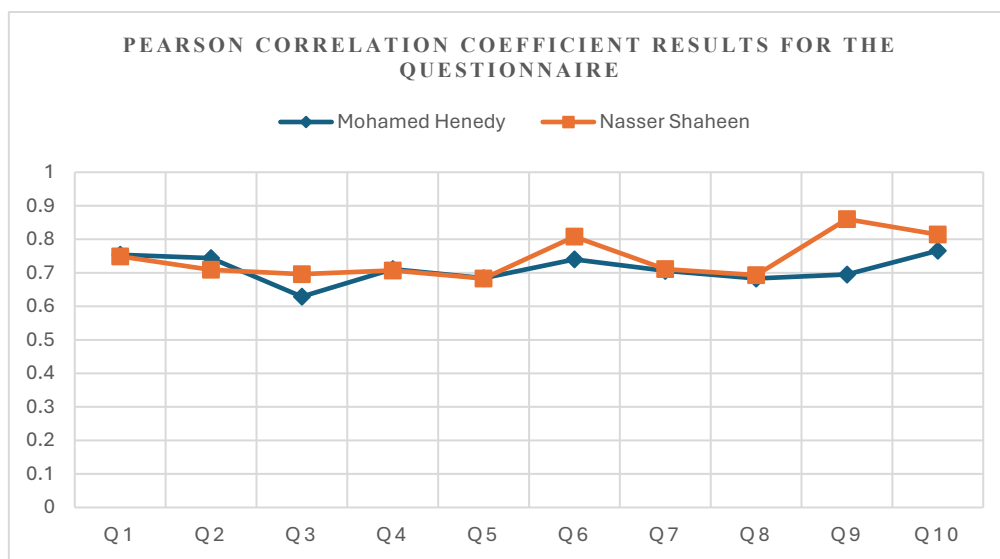


Figure 5. Pearson Correlation Coefficient Results of the Questionnaire

Comparing the correlation coefficients in Table 7, Mohamed Henedy displays stronger correlations between the specific tone elements like Energy (0.754), comedy (0.744), engagement (0.711), matching the character personality (0.684) and his overall effectiveness in the role of Timon. For Nasser Shaheen, certain factors like sarcasm (0.696), humor (0.808), enhancing the meaning of the character lines (0.711), enjoyability (0.693), and wit alignment (0.860) correlate more highly with his success, while other traits show weaker relationships.

The results of this study provide valuable insights that build on the existing literature regarding how vocal tone shapes audience reception of translated animated content, particularly for young viewers. The results align with several key strategies and priorities outlined by Chaume (2012) for successfully translating children's audiovisual media. Matching the tone and spirit of the original voice emerged as critical for authentically transporting the essence of Timon's characterization across cultures. Both Henedy and Shaheen utilized creative adaptation in their comedic delivery and sarcasm to resonate with young audiences, reflecting Chaume's (2012) emphasis on humor localization. Henedy's energetic and humorous dubbing specifically mirrors the conversational flow prioritization that Chaume (2012) advocates over literal accuracy.

Furthermore, the results reaffirm Perego's (2008) and Igareda's (2016) conclusions regarding comedic timing and musicality being integral to conveying the right tone for humorous animated characters. As the questionnaire responses showed, Henedy's vocal rhythms strongly conveyed Timon's witty comedic timing. This aligns with Perego's (2008) guidelines around timing preservation through word choice and inflection to transmit humorous personality. It also reflects Igareda's (2016) observations about matching comedic flow and musicality through expanded and localized dialogue.

The higher audience ratings for Henedy's dynamic comedic delivery despite looser fidelity also relate to O'Connell's (2003) and Martinez Sierra's (2004) points around balancing authenticity with comprehension and enjoyment. While Shaheen's portrayal had merits in conveying sarcasm, Henedy's more adaptive, engaging vocalization better balanced conveying Timon's essence while delighting young viewers. This showcases the nuanced tradeoffs between source accuracy and humor localization raised by O'Connell (2003) and Martinez Sierra (2004).

Furthermore, the questionnaire methodology and audience reception build on reception studies by González Vera (2018) and Igareda (2011), highlighting target culture voice acting preferences. The results revealing strong respondent preference for Henedy mirror Spanish audiences' dubbing inclinations found in those studies. This reemphasizes the vocal performance's role in driving enjoyable engagement and connection with characters, as Igareda's (2011) findings show. Reception-centered research provides vital insights into translation approaches that succeed with target audiences.

The in-depth examination of vocal tone also aligns with Maier's (2019) emphasis on considering multiple modalities together when adapting characterization. As Mera (1999) investigates music's narrative role, this study affirms that voice acting crosses verbal language to crucially shape personality and emotion. The results reinforce guidelines to holistically attend to dialogue, visuals, and audio elements to preserve a character's essence, as Maier suggests.

Although this analysis centered on vocal performance, Wikholm's (2012) views regarding risks of excessive domestication are worth considering regarding Timon's characterization in the Arabic translation overall. Further investigation could assess if Timon's representations relied on cultural stereotypes versus nuanced portrayal. As Marine-Roig (2015) notes, balanced translation fosters valuable intercultural learning even in children's media. Although this study focused specifically on vocal tone, broader cultural representation analysis would provide richer insights.

In summary, this study's focus on vocal performance in shaped animated character reception supports and expands on numerous facets of the literature around successfully translating children's media to engage young audiences across cultural contexts. The questionnaire methodology also provides direct audience feedback to complement existing theoretical discussions. Together these results underscore tone's integral role in building authentic cross-cultural connections between viewers and the essence of beloved animated characters.

V. CONCLUSION

This study explores how the vocal tone of different voice actors affects audience perception of an animated character in translation. By analyzing audience feedback on Mohamed Henedy and Nasser Shaheen's portrayals of Timon in the Arabic dubbing of *The Lion King's Timon and Pumbaa*, important insights emerge. The results show that Henedy's energetic, comedic, and engaging vocal style strongly aligned with Timon's characterization, making his performance more impactful and memorable. Although Shaheen succeeded in capturing Timon's wit and sarcasm, his portrayal was seen as slightly less well-rounded.

Ultimately, matching the unique tone and delivery of the original voice acting is crucial for transporting the essence of an animated character authentically in translation. Attention to vocal nuances can elevate a competent dubbing performance to truly iconic status in the minds of audiences. Based on the aforementioned, several promising directions for future research emerge. As Nicolás (2020) demonstrates with title translation analysis, paratextual elements beyond dialogue could be studied regarding their impact on engagement.

Exploring Arabic title translations for the *Timon and Pumbaa* series may reveal additional localization strategies. Furthermore, as González Vera (2018) notes, comparative reception studies on subtitling versus dubbing preferences could shed light on target culture norms. It would be interesting to compare responses with subtitled versions of the show to better understand Arabic audiences' expectations. Broader studies across multiple languages like Igareda's (2011) Spanish research could also illustrate wider trends and variations. Given the global popularity of *The Lion King* franchise, large-scale cross-cultural comparisons of Timon's translated receptions may highlight particularly impactful strategies. Additionally, characterizing viewer demographics, as Wissmath and Weibel (2012) do regarding age factors, could provide finer-grained insights. The questionnaire sample could be expanded and categorized by age to assess potential generational differences in translation style preferences.

Finally, follow-up discourse analysis, as Martínez Sierra (2004) models, could enrich the understanding of humor adaptation techniques. Based on this tonal study, closer examination of the linguistic and dialogical strategies used in translating Timon's wit could reveal underlying patterns. Combined with this project's vocal tone analysis, such multidimensional research directions can advance knowledge regarding how to create engaging, authentic translated experiences that allow iconic animated characters like Timon to delight children worldwide.

REFERENCES

- [1] Akan, F. M., Karim, R. M., & Chowdhury, K. A. (2019). An analysis of Arabic-English translation: Problems and prospects. *Advances in Language and Literary Studies*, 10(1), 58. DOI: 10.7575/aiac.all.v.10n.1p.58
- [2] Alsuhaime, D. (2024). Dubbing of English animated series into Arabic on Shahid and Netflix: an analysis based on the politeness theory. *Saudi Journal of Language Studies*, 4(2), 69-96. <https://doi.org/10.1108/SJLS-02-2024-0015>
- [3] Barbañán, S. G. (2019). The child's voice and perspective in Spanish dubbed versions of Pixar's Inside Out. In *Voice-over translation in the entertainment industry* (pp. 151-174). John Benjamins Publishing Company.
- [4] Bruti, S. (2009). The translation of compliments in subtitled cartoons. In J. Díaz-Cintas (Ed.), *New trends in audiovisual translation* (pp. 226-238). Multilingual Matters.
- [5] Bucaria, C., & Scifo, S. (2009). Humour, laughter and translation in children's books. *Translators' Journal*, 54(1), 186-196.
- [6] Chaume, F. (2021). Dubbing practices in Europe: Localisation beats globalisation. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 6. <https://doi.org/10.52034/lanstts.v6i.188>
- [7] Chaume, F. (2012). *Audiovisual translation: Dubbing*. St. Jerome Publishing.
- [8] González Vera, P. (2018). Translating for children in a context where dubbing prevails: Audience reception of audiovisual programmes in Spain. *Perspectives*, 26(5), 732-746.
- [9] Igareda, P. (2011). The use of voice-over in Disney and Pixar films in Spain. An empirical approach. *Trans*, 15, 127-134.
- [10] Igareda, P. (2016). Inside the songs: Translating Disney. *Perspectives*, 24(3), 442-458.
- [11] Lesnick, S. (2017). *Timon and Pumbaa Cast in Disney's Live Action Lion King!* Retrieved 22 June, 2022, from <https://www.comingsoon.net/movies/news/841623-timon-and-pumbaa>

- [12] Mazanec, T. J. (2011). *To know the tone: Analyses and experimental translations of Li Duan's poetic experiments* [Unpublished Master's thesis]. University of Colorado at Boulder.
- [13] Maier, C. D. (2019). The multimodality of animation: A model for analyzing speech, gesture, and graphic systems. *International Journal of Translation*, 31(1), 93-112.
- [14] Marine-Roig, E. (2015). Domestication and foreignization strategies in the translation of Disney comics. *Perspectives*, 23(1), 22-41.
- [15] Martínez Sierra, J. J. (2004). Estudio descriptivo y discursivo de la traducción del humor en textos audiovisuales. El caso de Los Simpson [Descriptive and discursive study of humor translation in audiovisual texts. The case of The Simpsons]. *Trans*, 8, 177-184.
- [16] Matamala, A. (2009). Main challenges in the translation of documentaries. In J. Díaz-Cintas (Ed.), *New trends in audiovisual translation* (pp. 109-120). Multilingual Matters.
- [17] Mera, M. (1999). Read my lips: Re-evaluating subtitling and dubbing in Europe. *Links & Letters*, 6, 75-87.
- [18] Mohammadi, L. (2023). Ta'thīr al-dablah bi-al-gharīb 'alā talaqqī al-atfāl al-Jazā'irīyīn li-al-rusūm al-mutaḥarrikah [تأثير الدبلجة بالتغريب على تلقي الأطفال الجزائريين للرسوم المتحركة]. *Majallat fī al-Tarjamah*, (2), 184-197. <https://asjp.cerist.dz/en/article/234316>
- [19] Nicolás, C. (2020). The translation of Disney film titles into Spanish. A descriptive study. *Sendebare*, 31, 257-284.
- [20] O'Hagan, M. (2009). Evolution of user-generated translation: Fansubs, translation hacking and crowdsourcing. *Journal of Internationalization and Localization*, 1(1), 94-121.
- [21] O'Connell, E. (2003). What dubbers of children's television programmes can learn from translators of printed literature. *Meta: Journal des Traducteurs Meta:/Translators' Journal*, 48(1-2), 222-232.
- [22] Perego, E. (2008). Subtitles and line-breaks: Towards improved readability. In D. Chiaro, C. Heiss & C. Bucaria (Eds.), *Between Text and Image: Updating research in screen translation* (pp. 211-223). John Benjamins Publishing Company. <https://doi.org/10.1075/btl.78.21per>
- [23] Perez-Gonzalez, L. (2014). *Audiovisual translation: Theories, methods and issues*. Routledge.
- [24] Rowe, R. (2018). Accentedness in Disney's Zootopia: The Other redux. In *Audiovisual translation in applied linguistics* (pp. 141-162). John Benjamins Publishing Company.
- [25] Saed, H., Haider, A. S., Tair, S. A., & Darwish, N. (2024). Dubbing English edutainment: Localizing educational content for Arab children. *Research Journal in Advanced Humanities*, 5(3), 391-406. <https://doi.org/10.58256/bh9fhj11>
- [26] Saiddina, D. S., & Darma, V. P. (2024). A study of early childhood English language development as an impact of English cartoons. *Journal of Research on English and Language Learning (J-REaLL)*, 5(1), 123-134. <https://doi.org/10.33474/j-reall.v5i1.21302>
- [27] Suchet, M. (2013). *Voice, Tone and Ethos: A Portrait of the Translator as a Spokesperson*. Retrieved 22 June, 2022, from <https://core.ac.uk/download/pdf/10990459.pdf>
- [28] Wang, Q., Noordin, N., Joseph Jeyaraj, J., & Sun, Z. (2024). Learners' perception towards dubbing on oral language learning: A systematic review (2017-2023). *International Journal of Academic Research in Business and Social Sciences*, 14(6), 799-816. <https://doi.org/10.6007/IJARBS/v14-i6/21690>
- [29] Wikholm, A. (2012). *Microrhetorics in dubbing and subtitling* [Doctoral dissertation]. Indiana University.
- [30] Wissmath, B., & Weibel, D. (2012). Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment. *Journal of Media Psychology*, 24(3), 113-125.

Shaymaa Y. S. Alharbi, an Assistant Professor at Taibah University. She holds a Ph.D. in Translation from the University of Leicester, United Kingdom (2019). She has over ten years of academic and teaching experience in English language and translation. She has participated in numerous training programs and academic conferences both locally and internationally, delivered workshops on translation fundamentals, and obtained a Train the Trainer (TOT) certification as well as executive diplomas in management, leadership, and total quality management.

Exploring Students' Perspectives on Quizlet as a Gamified Vocabulary Learning Tool in a Business English Course

Thuan D. Pham

Department of Foreign Languages and Information Technology, Hoa Lu University, Ninh Binh, Vietnam

Tam T. Pham

Department of Foreign Languages, Academy of Finance, Ha Noi, Vietnam

Abstract—This study explores university students' perspectives on Quizlet as a gamified vocabulary learning tool in a business English course. Conducted with fifteen third-year business administration students at a university in Vietnam, the study employed focus group interviews as a data collection tool to examine students' perceptions of Quizlet. Findings reveal that students overwhelmingly found Quizlet user-friendly, accessible, and engaging, with its gamified features enhancing motivation and vocabulary retention. The platform's various study modes, particularly flashcards, quizzes, and Quizlet Live, contributed to increased active participation and long-term recall. While previous studies caution against over-reliance on gamified learning for rote memorization, this research demonstrates that Quizlet facilitates both memorization and contextual learning when strategically integrated into instruction. The study contributes to the growing body of research on technology-enhanced language learning, particularly in English for Specific Purposes (ESP) vocabulary acquisition. Despite its positive findings, the study acknowledges limitations, including its small sample size and qualitative focus, suggesting avenues for future research such as longitudinal studies on retention and comparative analyses of gamified tools in ESP contexts. Findings offer practical insights for educators seeking to optimize digital learning strategies for business English vocabulary instruction.

Index Terms—business English, gamified learning, Quizlet, students' perceptions, vocabulary learning

I. INTRODUCTION

Vocabulary acquisition stands as a cornerstone in ESP courses, particularly when focusing on business vocabulary (Coxhead, 2012; Hrdličková, 2020; Normurodovna, 2025; Panigrahi, 2014). As Brooks (2016) highlights, vocabulary plays a crucial role in ESP teaching and learning, especially within Asian university contexts, emphasizing the necessity of targeted vocabulary instruction to achieve communicative competence. In business English, the acquisition of specialized terms is not merely about memorization; it's about understanding and applying these terms within professional contexts (Budinić, 2022). One of the central challenges learners face is the sheer volume and specificity of business vocabulary necessitating a focused approach to vocabulary selection and teaching. Tangpijaikul (2014) and Hrdličková (2020) further elaborate on this, discussing the need for careful preparation of business vocabulary for the ESP classroom, suggesting that effective vocabulary instruction must be tailored to the specific needs of the learners and the demands of their professional fields. Chirobocea (2018) delves into the perspectives, strategies, and resources for vocabulary acquisition in ESP, advocating for a multifaceted approach that includes various learning strategies and resources to enhance vocabulary retention and application. Moreover, research has shown that traditional vocabulary learning methods often fail to maintain student interest and engagement (Tran & Le, 2023; Václavík, 2020).

The integration of technology into language learning has transformed traditional pedagogical approaches, introducing innovative gamified learning tools aimed at enhancing learners' engagement and efficiency. Among these tools, Quizlet has emerged as a popular gamified learning platform, offering various interactive features designed to facilitate vocabulary retention and comprehension (Dizon, 2016; Fox, 2020). Quizlet employs digital flashcards, games, and quizzes to create an engaging learning environment, which can foster motivation and active participation (Sanosi, 2018; Wright, 2016). Prior research has demonstrated the platform's effectiveness in promoting vocabulary acquisition across different educational contexts (Al-Malki, 2020; Anjaniputra & Salsabila, 2018). However, despite its widespread use, there remains a lack of comprehensive research on university students' perspectives regarding the use of Quizlet specifically in business English courses. Understanding these perspectives is essential to optimizing the integration of gamified tools in language curricula.

Given the increasing emphasis on business communication skills in professional settings, understanding how students perceive Quizlet's role in acquiring industry-specific terminology is crucial. This study addresses this gap by investigating university students' experiences with Quizlet in a business English course. The research question guiding the study is: "How do students perceive Quizlet as a vocabulary learning tool in a business English course?" The

rationale for conducting this study lies in the need to optimize digital learning strategies for ESP courses, ensuring that gamified tools like Quizlet are effectively integrated into business English instruction. By exploring students' perspectives, this study aims to provide educators with insights into the pedagogical value of Quizlet, informing future curriculum design and instructional practices to enhance business vocabulary learning.

II. LITERATURE REVIEW

A. Gamification in Language Learning

Gamification, defined as the use of game elements in non-game contexts, has increasingly been integrated into language education to boost student motivation and engagement (Nilubol & Sitthitikul, 2025; Zhang & Hasim, 2023). Initially popularized through digital platforms, gamification strategies have evolved from basic reward systems to more sophisticated tools incorporating narratives, challenges, and competitive elements (Hardianti, 2024). Recent studies suggest that the incorporation of gamified elements aligns with contemporary pedagogical theories such as self-determination theory, which emphasizes intrinsic motivation (Liu et al., 2024). The growing accessibility of digital tools has also accelerated the adoption of gamification, particularly in online and hybrid learning environments (Nilubol, 2023).

In English as a Foreign Language (EFL) contexts, gamified learning tools like Quizlet, Quizizz, and Kahoot have been employed to enhance various language skills, including vocabulary, grammar, and writing (Jiménez-Sánchez & Gargallo-Camarillas, 2020; Panmei & Waluyo, 2022). Ardi and Rianita (2016) demonstrated that integrating gamification into grammar lessons led to increased student engagement and participation. Similarly, Pingmuang and Koraneekij (2022) found that mobile-assisted gamification strategies effectively enhanced EFL learners' writing skills through interactive, task-based activities. In large EFL classes, gamified tools have been particularly effective in maintaining student attention and fostering collaborative learning (Kraus et al., 2020).

Gamification offers several pedagogical benefits in language education including enhancement of motivation and engagement (Rofiah & Waluyo, 2024; Sadeghi et al., 2022), development of learner autonomy (Rahmani, 2020; Zhou, 2024), improvement of academic performance in EFL learners (Tayeh et al., 2024), promotion of collaborative learning (Redjeki & Muhajir, 2021) and stimulation of vocabulary retention by offering varied practice modes and creating meaningful learning experiences (Ahmed et al., 2022; Ika, 2024; Setiawan & Wiedarti, 2020; Zhang & Crawford, 2023; Waluyo & Bucol, 2021). Despite these advantages, Phuong (2020) cautions that technological literacy is a critical factor influencing the effectiveness of gamification in different educational contexts.

B. Quizlet as a Vocabulary Learning Tool

Quizlet is a popular online platform that provides a variety of learning tools, including flashcards, games, and quizzes, for various subjects, including language learning (Al-Malki, 2020; Anjaniputra & Salsabila, 2018; Fox, 2020); and cater to various learning styles (Dizon, 2016; Fox, 2020). Its user-friendly interface and accessibility across multiple devices make it a versatile tool for both in-class and out-of-class learning (Luu & Doan, 2021; Vogel, 2022). Quizlet allows users to create their own study sets or utilize pre-existing ones, catering to diverse learning needs and preferences (Bueno-Alastuey & Nemeth, 2020; Dine, 2018).

The platform offers various study modes, such as "Learn," "Write," "Spell," "Test," and "Match," which cater to different learning styles and promote active recall and spaced repetition, essential strategies for effective vocabulary acquisition (Ho & Kawaguchi, 2021; Platzer, 2020). Furthermore, Quizlet's gamified features, such as "Quizlet Live," can foster collaborative learning and friendly competition in the classroom (Wright, 2016). The platform's ability to track progress and provide feedback allows learners to monitor their learning and identify areas for improvement (Boroughani et al., 2023). Several studies have consistently affirmed Quizlet's effectiveness in enhancing vocabulary acquisition across a range of EFL and ESP contexts (Al-Malki, 2020; Chaikowska & Zbaravska, 2020; Dewi & Rahma, 2022; Dizon, 2016; Hamida et al., 2023; Ho et al., 2020; Ho & Kawaguchi, 2021; Lubis et al., 2022; Mykytka, 2023; Nur & Putri, 2023; Sanosi, 2018; Setiawan & Wiedarti, 2020; Todorova, 2023; Tran & Le, 2023; Tran & Nguyen, 2022; Wulandari & Irfani, 2023). These studies highlight not only improvements in retention and recall but also demonstrate how Quizlet's interactive, gamified features foster learner autonomy and engagement, especially when integrated consistently into formal instruction or self-directed study routines. Furthermore, the diversity of learner populations and instructional settings in this body of research underscores Quizlet's adaptability and pedagogical relevance across cultures and disciplines, suggesting its strong potential for wide-scale implementation in vocabulary learning. Research indicates that Quizlet effectively enhances both receptive and productive vocabulary knowledge (Ho & Kawaguchi, 2021; Mohammadi et al., 2024). According to Al-Malki (2020) and Tran and Nguyen (2022), Quizlet's interactive features facilitate active engagement, which contributes to improved vocabulary retention. Moreover, studies have demonstrated that students perceive Quizlet as an easy-to-use platform, which positively influences their willingness to adopt it for language learning (Aprilani & Suryaman, 2021; Lukov, 2022).

C. Previous Studies on Students' Perspectives on Quizlet

A growing body of research has explored students' perceptions of Quizlet as a vocabulary learning tool. Many studies have reported positive student attitudes towards Quizlet, highlighting its user-friendliness, engaging features, and

effectiveness in improving vocabulary knowledge (Aprilani & Suryaman, 2021; Bayaksud et al., 2024; Nhac & Ngo, 2023; Ningsih et al., 2024; Tran & Nguyen, 2023; Trinh & Ho, 2023). Students often appreciate the flexibility and accessibility of Quizlet, allowing them to study anytime and anywhere (Ho et al., 2020; Luu & Doan, 2021; Zakian et al., 2022). The various study modes offered by Quizlet are also valued by students, as they cater to different learning preferences and promote active learning (Anjaniputra & Salsabila, 2018). Moreover, the gamified elements of Quizlet, such as "Quizlet Live," are often seen as motivating and enjoyable, contributing to increased engagement and participation (Montaner-Villalba, 2019; Sangtuptim, 2018).

Research highlighted the role of Quizlet in facilitating independent learning, with students appreciating the self-paced nature of the platform (Al-Malki, 2020; Nguyen & Nguyen, 2021; Nguyen et al., 2021). Likewise, Aprilani and Suryaman (2021) found that learners using Quizlet experienced improved vocabulary retention, attributing their success to the platform's repetitive learning features and personalized study modes. Mykytko (2023) observed that students appreciated Quizlet's accessibility and ease of use, which allowed for consistent vocabulary practice outside the classroom. A study conducted by Dizon (2016) with Japanese university students revealed a significant improvement in academic vocabulary acquisition through regular use of Quizlet. In the same way, Ho and Kawaguchi (2021) demonstrated that EFL learners' receptive vocabulary acquisition significantly improved due to Quizlet's interactive features. Findings from Tran and Le (2023) further emphasized the positive influence of Quizlet on students' confidence and motivation in learning business English vocabulary. Moreover, Rofiah and Waluyo (2024) examined the effects of gamified grammar and vocabulary learning on EFL students in Thailand, concluding that Quizlet significantly increased both engagement and motivation. Özdemir and Seçkin's (2024) meta-analysis also confirmed the cognitive and affective benefits of using Quizlet, particularly in terms of improving learning outcomes and fostering a positive learning environment.

In the context of specialized vocabulary, Nhac and Ngo (2023) explored students' use of Quizlet for enhancing legal English vocabulary. The study found that learners perceived the platform as particularly beneficial for retaining complex terminology and fostering long-term vocabulary retention. Ningsih et al. (2024) observed a positive impact on the transferability of vocabulary knowledge when learners engaged with Quizlet's various modes, such as testing and matching games.

Despite its many benefits, some studies highlighted potential drawbacks. According to Lubis et al. (2022), students faced difficulties with technical issues and limited customization of content. Zakian et al. (2022) noted concerns over the platform fostering rote memorization rather than deep learning, particularly when students focused excessively on repetitive drills rather than context-based usage. Nur and Putri (2023) pointed out that students occasionally relied too heavily on Quizlet, neglecting other valuable learning methods. Analogously, Sadeghi et al. (2022) cautioned against overemphasizing gamification elements, as it could lead to superficial engagement with learning materials.

III. METHODOLOGY

A. Context and Participants

This research was conducted in an ESP course during the second semester of the 2023-2024 academic year at a university in northern Vietnam. The institution provides a range of majors including Culture and Tourism, Accounting, Business Administration, Information Technology, Primary Education, Preschool Education, Mathematics Education, History and Geography Education, and Natural Science Education. At this university, English is taught as a foreign language and is a mandatory course for students across every major. English courses include standard English and ESP. General English comprises three courses: English 1, English 2, and English 3. Each academic discipline often offers an ESP course tailored to that field of study. Students are required to finish the standard English courses prior to engaging in ESP. This research included 15 students specializing in Business Administration. Business administration students are required to complete three specific English courses: Basic Business English 1, Basic Business English 2, and Basic Business English 3.

Fifteen students included in this research were third-year undergraduates, including six males and nine females. They were non-English majored, whose English proficiency level was estimated at around pre-intermediate level. They completed general English courses and Basic Business English 1 and 2. They studied Basic Business English 3 using the New Market Leader Intermediate textbook series produced by Longman. This textbook consists of 14 modules, each addressing distinct themes, and follows a consistent structure: Starting Up, Vocabulary, Discussion, Reading, Listening, Language Review, Skills, and Case Study. The course comprises 4 credits and spans 15 weeks, with two in-person classroom sessions each week.

B. Application of Quizlet in the Learning Process

The Quizlet implementation process lasted 8 weeks, from week 7 to 14 in the syllabus. First, the class was created on the platform, and students were sent the link to join the class by their emails. And at the end of each live session in class, students were invited to come to the board and write down 10 new words related to the topic of the lesson that they needed to practice more. These words were then generated in Quizlet in different sets. In total, 14 new word sets were

created covering themes of employment, trade, quality, ethics, leadership, innovation, and competition. In each vocabulary set, new words were presented along with definitions in Vietnamese and a picture as an illustration.

The vocabulary sets were used in 2 ways. First, students were sent a link to practice outside of class time. The levels and progress of the students were saved by the system and announced to all students in the next class session so that students could grasp their results. Second, in the last 15 minutes of each class, to review the word sets, students had to spend about 5 minutes looking at the words again and spend 10 minutes playing Quizlet Live all together before the end of the class. To play Quizlet Live, QR codes were shown along with pass codes of the game on the projector screen in the classroom for the students to scan to log in to the game using their smartphones.

C. Data Collection and Analysis

The qualitative approach, which prioritized process, meaning, and comprehension, was applied in this investigation (Flick, 2023; Merriam, 2009). The researcher served as the main tool for gathering and analyzing data. Data for the research was gathered via focus group interviews. Near the conclusion of the course, the students were invited to attend interviews. All 15 students agreed to participate in the interviews voluntarily. To be comfortable and convenient, the students were encouraged to express their opinions throughout the interviews in their native language Vietnamese. The question for the participants was: "What do you think about Quizlet as it was integrated in the learning process?" The interviews, which took place in the teacher lounge next to the classroom, lasted for about ninety minutes and audio recorded. All of the students were encouraged to freely share their thoughts of Quizlet as it was integrated in the course as a tool to support the learning new vocabulary items. Later, English transcriptions and translations were made of the recordings.

Because it offered a methodical and objective approach to drawing reliable conclusions from spoken, observed, or written data, content analysis was selected as the data analysis technique. This made it simpler to identify themes and patterns in the data (Creswell, 2014). Because the goal of this research was to identify students' perspectives on Quizlet as a gamified vocabulary learning tool in learning business English - a task that can only be accomplished with qualitative, descriptive data - this approach worked well. The procedure had several steps: getting familiar with the collected data through repeated reading, organizing text passages, interpreting and classifying the coded labels into themes found in the literature study. This method was chosen due to its capacity to provide comprehensive qualitative insights into students' experiences, demonstrating how they perceived Quizlet in the process of learning new business English words.

IV. RESULTS

The results analyzed from the gathered data show that students overwhelmingly found Quizlet to be an exceptionally user-friendly tool for learning business vocabulary, emphasizing its intuitive design and ease of navigation, engagement, flexibility and accessibility, and effectiveness in reinforcing vocabulary retention. Many students mentioned that they could log in and start using the platform without any prior guidance, as the interface was clean, well-organized, and simple to understand. Student 3 stated: *"It's pretty straightforward. I mean, you don't need a manual to figure it out, just pick a set and go."* In parallel, Student 5 noted: *"I thought it'd be complicated, but it's not. The interface is so intuitive. I just hit 'play' and start matching words like 'dividend' to definitions"*.

Unlike traditional vocabulary learning methods that may involve rote memorization, Quizlet's structured yet effortless layout allows students to focus on the learning process rather than the tool itself. The ability to quickly access pre-made decks or create personalized vocabulary sets was particularly appreciated, as Student 4 remarked: *"The app's design is clean, so I don't get lost. I can make my own business English decks or use someone else's in like two seconds. It's a time-saver."* The smooth functionality of flashcards, quizzes, and interactive exercises made the learning experience efficient, even for those who did not consider themselves tech-savvy. This ease of access and usability fosters autonomy and self-directed learning, enabling students to engage with vocabulary at their own pace and convenience.

Another significant aspect that students highlighted was the engaging, gamified nature of Quizlet's learning activities, which transformed vocabulary acquisition into a more enjoyable experience. Many students felt that the platform did not feel like studying at all, as the interactive features – such as matching games, quizzes, and progress tracking - helped maintain their motivation and interest in learning. Student 1 described this aspect, saying: *"Quizlet's super easy to use. I just log in, find my business vocab set, and start playing the matching game. It doesn't feel like studying at all, which is awesome."* Likewise, Student 7 stated: *"It's almost too easy to use, honestly. I just search for business vocab sets, hit the game mode, and I'm learning 'stakeholder' while having fun."* The game-like approach to studying not only made learning less monotonous but also helped with better retention of business vocabulary terms such as "leverage," "dividend," and "acquisition." By incorporating elements of play, Quizlet provided a sense of achievement, where students could monitor their own progress and feel a rewarding sense of accomplishment when mastering new words. Student 6 highlighted the motivational aspect, explaining, *"I can study on my phone anywhere, and the little progress bar keeps me motivated to nail terms like 'acquisition' without stress."* This aligns with research on gamified learning, which suggests that engagement and motivation are crucial in enhancing knowledge retention and recall.

The flexibility and accessibility of Quizlet also contributed to its effectiveness as a vocabulary learning tool. Students appreciated the fact that they could study anytime and anywhere, using either mobile devices or computers, which allowed for seamless integration into their daily routines. Student 11 supported this, saying: *"Quizlet's dead simple. I can pull up a business English deck on my laptop or phone and test myself on 'market share' in like five minutes flat"*.

The ability to study in short, focused sessions rather than lengthy, structured lessons made Quizlet an adaptable and convenient tool for learning business vocabulary in small, manageable increments. Student 9 noted this convenience, stating: *"I love how quick it is to get started. You don't need to mess around; just pick a set, play, and suddenly 'cash flow' sticks in your head"*.

This flexibility is particularly beneficial for busy university students, as it enables them to review and reinforce vocabulary on the go, whether during short breaks, commutes, or between classes. The availability of various study modes meant that learners could switch between different approaches based on their preferences or needs, further enhancing the learning experience. Finally, students noted that Quizlet was effective in reinforcing vocabulary retention, as the repetitive and interactive nature of the platform made it easier to remember words. The multiple learning modes, including flashcards, quizzes, and matching activities, provided diverse exposure to the same vocabulary items, which strengthened long-term memory recall. Student 14 reflected on this, saying: *"It's honestly effortless. I load up a set, swipe through some flashcards, or play the gravity game, and words like 'outsourcing' just sink in"*.

Unlike passive learning techniques, Quizlet's active recall mechanisms, such as self-testing and spaced repetition, helped students internalize words more efficiently. Many students expressed that business terms like "benchmarking," "collaboration," and "negotiation" became easier to remember because of their repeated exposure in different engaging formats. Student 12 explained this effectiveness, stating: *"The ease of use is unreal. I just tap around, play a round or two, and boom. I've got 'collaboration' locked in my brain. No fuss"*.

This suggests that Quizlet's structure aligns well with cognitive learning theories, which emphasize active engagement and multimodal exposure as key factors in successful language acquisition. Obviously, students' perspectives on Quizlet as a gamified business vocabulary learning tool were highly positive, with ease of use, engagement, flexibility, and effectiveness emerging as the dominant themes. The tool's ability to make vocabulary learning simple yet stimulating, combined with its on-the-go accessibility, makes it a powerful and efficient resource for EFL students in an ESP context. By turning vocabulary acquisition into an interactive and rewarding experience, Quizlet not only improves learning outcomes but also fosters a more enjoyable and self-motivated approach to studying business English. As Student 15 summed up: *"Using Quizlet feels natural. I don't even think about the app - just the vocab. I can master 'sustainability' while chilling on the couch."* This underscores the platform's ability to blend learning seamlessly into students' daily lives, making it an effective and engaging educational tool.

Despite the overall positive attitudes towards Quizlet, several students expressed challenges in retaining newly learned vocabulary over time. Student 3 noted:

I like learning vocabulary with Quizlet. But I found that sometimes I could not remember the meanings of difficult words after several days when I did the reading comprehension tasks in the textbook. Therefore, I need to review the words regularly.

Additionally, Student 5 observed a short-term gain followed by forgetfulness: *"I can remember the new words right away after Quizlet games in class. However, the next days I checked again I forgot some words."* Student 6 also reflected on personal struggles with retention: *"Although I practice at home with Quizlet, I forget the words when I join the review game in class and check the word meaning. I often rank after many other classmates."* These comments indicate that while Quizlet may support initial acquisition and in-class engagement, sustaining long-term vocabulary retention remains a concern for some learners, particularly those with weaker memory or inconsistent review habits. These findings emphasize the need for structured spaced repetition and reinforcement strategies when using digital tools like Quizlet.

V. DISCUSSION

A. Perceived Usefulness and Learning Effectiveness

Students consistently described Quizlet as helping them retain and recall business vocabulary more efficiently, confirming the tool's established role in promoting vocabulary learning. The results echo empirical findings demonstrating the effectiveness of digital flashcards and spaced repetition. This aligns strongly with results from Dizon (2016), Sanosi (2018), Dewi and Rahma (2022), and Ho and Kawaguchi (2021), all of whom reported significant improvement in vocabulary test performance among EFL learners using Quizlet. By the same token, Montaner-Villalba (2019) and Mykytka (2023) found that Quizlet promotes frequent exposure and retrieval practice – key elements in long-term vocabulary retention. In the context of ESP learning, the results resonate with Václavík (2020), Hrdličková (2020), and Chirobocea (2018), who emphasize that technology-mediated vocabulary review supports learners in managing the large quantity of technical terminology required in business English. Several students in the present study explicitly cited Quizlet as helping them "keep track of business vocabulary" or "organize terms systematically," illustrating how digital tools reduce cognitive load when learning field-specific terminology. However, the current study expands earlier findings by showing that students' perceived usefulness is tightly tied to controlled, repeated interaction

with vocabulary rather than deeper contextual understanding. While prior research widely affirms the gains in receptive vocabulary (e.g., Ho & Kawaguchi, 2021; Tran & Nguyen, 2022), this study's qualitative data highlights a gap between memorization and functional use, suggesting the need for integrated instructional strategies that bridge Quizlet practice with context-based ESP tasks.

B. Motivational and Affective Engagement through Gamification

Learners repeatedly highlighted Quizlet's gamified features – such as Match, Gravity, flashcards, and progress tracking – as motivating and enjoyable. Students described learning as “fun,” “less stressful,” and “like playing a game,” reflecting the same affective benefits reported in prior research on digital gamification (Liu et al., 2024; Sadeghi et al., 2022; Redjeki & Muhajir, 2021). These results also align with Ahmed et al. (2022), Rahmani (2020), and Panmei and Waluyo (2022), who found that game-based elements significantly increase engagement and persistence among EFL learners. Additionally, consistent with Luu and Doan (2021) and Waluyo and Bucol (2021), the present study shows that gamification fosters not only enjoyment but also sustained, self-directed engagement, especially when learners can access Quizlet outside class. The findings extend the literature by revealing dual affective responses: while gamification enhances enjoyment, some students expressed concerns about being “too focused on winning the game” rather than learning deeply. This echoes the caution raised by Wright (2016), who argued that gamified learning may risk superficial processing if instructional design does not integrate reflective or analytic tasks. Thus, the present study affirms that gamification can energize vocabulary learning but must be balanced with strategies supporting deeper lexical understanding, especially in ESP contexts.

C. Ease of Use, Digital Accessibility, and Autonomous Learning

Students overwhelmingly praised Quizlet's accessibility, clean interface, and cross-platform compatibility, describing it as “easy,” “smooth,” and “simple to navigate.” This supports earlier findings by Aprilani and Suryaman (2021), Al-Malki (2020), and Lubis et al. (2022), who reported that usability is one of Quizlet's central strengths that increases learner adoption. In line with Nguyen et al. (2021) and Nguyen and Nguyen (2021), the present study shows that ease of use directly enhances learner autonomy. Many participants reported using Quizlet independently at home, while traveling, or between classes—behaviors indicating high self-regulatory engagement, which is also emphasized in Boroughani et al. (2023) and Tran and Nguyen (2022). These findings further reinforce the role of mobile-assisted vocabulary learning highlighted in studies such as Zakian et al. (2022) and Mykytka (2023), which demonstrate that mobile digital flashcards promote continuous learning beyond classroom boundaries. In the ESP domain, compatibility with busy student schedules is crucial, and our results support the argument of Brooks (2016) and Budinčić (2022) that flexible vocabulary tools are essential for supporting domain-specific language development. The present study deepens the discussion by showing that ease of use is not merely a technical benefit, but a pedagogical enabler, allowing students to engage in short, repeated, self-directed learning episodes—essential for mastering complex business vocabulary.

D. Challenges and Limitations of Quizlet for ESP Vocabulary Learning

Although the findings affirm Quizlet's many strengths, students also identified limitations, offering a more nuanced picture than many prior studies. Several participants expressed that Quizlet helps with memorization but not with deeper understanding, particularly regarding business concepts requiring contextual or situational interpretation. This critique supports concerns raised in the ESP literature (Panigrahi, 2014; Tangpijaikul, 2014) that rote vocabulary practice is insufficient for mastering the conceptual depth needed in business English. While Václavík (2020) and Wulandari and Irfani (2023) noted Quizlet's usefulness for business terminology review, they also emphasized the need for teacher scaffolding and authentic tasks – findings echoed in this study. Furthermore, learners mentioned the risk of “overreliance on match games” and distraction, aligning with Wright's (2016) argument that gamification may sometimes shift focus from meaning to speed. This adds nuance to highly positive results in earlier quantitative studies such as Dewi and Rahma (2022) and Hamida et al. (2023), suggesting that effectiveness depends on how Quizlet is integrated into pedagogy rather than on the tool alone. Thus, the present study contributes a critical perspective that, while Quizlet is effective for initial exposure and review, ESP instructors must supplement it with explanation of business concepts, contextualized reading and case-based tasks, collaborative problem solving, and discussion-based application. By doing so, vocabulary learning supports professional communication needs rather than solely receptive recall.

VI. CONCLUSION

This qualitative study explored Vietnamese university students' perspectives on using Quizlet as a gamified vocabulary learning tool in a Business English course. The results reveal that students generally hold highly favorable attitudes towards Quizlet, particularly emphasizing its intuitive design, ease of access, and motivational value. The gamified features such as flashcards, match games, and learning modes were repeatedly mentioned as enhancing their learning experience, reducing monotony, and improving vocabulary retention. These insights confirm and deepen previous findings on gamified mobile-assisted language learning tools, affirming Quizlet's role as an effective, student-friendly platform in ESP contexts. The study contributes to existing literature by offering nuanced, context-specific

insights into how students engage with vocabulary learning through digital gamification. While much of the earlier research has emphasized Quizlet's general benefits in EFL or general English courses, this study enriches the field by focusing on its application in Business English vocabulary learning—a relatively underexplored area. The findings add to the growing body of evidence supporting gamification's potential in increasing learner motivation, autonomy, and long-term vocabulary retention.

Pedagogically, this study suggests that incorporating gamified platforms like Quizlet into ESP instruction can foster learner engagement and enhance vocabulary acquisition. Instructors are encouraged to curate or co-create relevant vocabulary sets tailored to course objectives and student needs. The flexibility and accessibility of Quizlet also make it an ideal complement to both classroom-based and self-directed learning environments. Moreover, Quizlet's interactive features support differentiated instruction and can be especially beneficial in large, mixed-ability classrooms.

However, the study has several limitations. First, it was conducted with a relatively small sample of students at a single Vietnamese university, which may restrict the generalizability of the findings. Second, the focus on qualitative data limits the ability to make statistical inferences regarding the effectiveness of Quizlet on vocabulary test performance. Third, although the study focused on student perceptions, triangulating these findings with teacher perspectives or classroom observation data could have provided a more holistic view of the learning process. Finally, since students used pre-made Quizlet sets in most cases, the study did not explore the full impact of student-generated content or collaborative vocabulary building. Future research should expand the scope to include diverse institutional settings, broader ESP areas beyond business, and mixed-method or longitudinal designs to measure learning gains over time. Comparative studies between Quizlet and other gamified tools (e.g., Kahoot!, Quizizz) could also yield insights into which features are most effective for vocabulary learning. Additionally, investigating the pedagogical strategies instructors employ when integrating Quizlet into their curricula would help clarify how to maximize the platform's benefits in practice.

In conclusion, this study underscores the positive perceptions Vietnamese university students hold towards Quizlet as a gamified vocabulary learning tool in Business English courses. By illuminating learners' experiences, this research supports the continued integration of digital gamified tools into ESP instruction, while also calling for more targeted and expansive future inquiries into their pedagogical affordances.

REFERENCES

- [1] Ahmed, A. A. A., Widodo, M., Komariah, A., Hassan, I., Sukmana, N., Ali, M. H., Abbas, A. K., & Rohi, A. (2022). Assessing the effects of gamification on developing EFL learners' idiomatic knowledge: Do attitudinal factors contribute to the learning of the idioms with the game? *Education Research International*, 2022, 1–10. <https://doi.org/10.1155/2022/2482570>
- [2] Al-Malki, M. A. (2020). Quizlet: An online application to enhance EFL foundation students' Vocabulary acquisition at Rustaq College of Education, Oman. *Arab World English Journal*, 6(2020), 332–343. <https://doi.org/10.24093/awej/call6.22>
- [3] Anjaniputra, A. G., & Salsabila, V. A. (2018). The merits of Quizlet for vocabulary learning at tertiary level. *Indonesian EFL Journal*, 4(2), 1–11. <https://doi.org/10.25134/ieflj.v4i2.1370>
- [4] Aprilani, D. N., & Suryaman, M. (2021). Students' perception in learning English vocabulary through Quizlet. *JET (Journal of English Teaching)*, 7(3), 343–353. <https://doi.org/10.33541/jet.v7i3.3064>
- [5] Ardi, P., & Rianita, E. (2016). Leveraging gamification into EFL grammar class to boost student engagement. *The Journal of Teaching English with Technology*, 22(2), 90–114. Retrieved February 19, 2025, from <https://tewtjournal.org/download/leveraging-gamification-into-efl-grammar-class-to-boost-student-engagement-by-priyatno-ardi-and-elvira-rianita/>
- [6] Bayaksud, N., Degeng, P. D. D., & Razali, K. A. (2024). The use of Quizlet for vocabulary learning: A lesson from innovative application. *Journal of Language and Literature Studies*, 4(1), 244–255. <https://doi.org/10.36312/jolls.v4i1.1823>
- [7] Boroughani, T., Behshad, N., & Xodabande, I. (2023). Mobile-assisted academic vocabulary learning with digital flashcards: Exploring the impacts on university students' self-regulatory capacity. *Frontiers in Psychology*, 14(2023). <https://doi.org/10.3389/fpsyg.2023.1112429>
- [8] Brooks, M. (2016). The role of vocabulary in English for specific purpose (ESP) teaching and learning: Considerations for Asia university. *CELE Journal*, 22, 157–175. Retrieved April 15, 2025, from <https://asia-u.repo.nii.ac.jp/records/25355>
- [9] Budinčić, V. (2022). Theoretical and practical aspects of teaching business English vocabulary in EFL classes. In *Topical Issues of Linguistics and Teaching Methods in Business and Professional Communication - TILTM 2022*. <https://doi.org/10.15405/epes.22104.7>
- [10] Bueno-Alastuey, M. C., & Nemeth, K. (2020). Quizlet and podcasts: effects on vocabulary acquisition. *Computer Assisted Language Learning*, 35(7), 1–30. <https://doi.org/10.1080/09588221.2020.1802601>
- [11] Chaikovska, O., & Zbaravska, L. (2020). The efficiency of Quizlet-based EFL vocabulary learning in preparing undergraduates for state English exam. *Advanced Education*, 7(14), 84–90. <https://doi.org/10.20535/2410-8286.197808>
- [12] Chirobocea, O. (2018). Vocabulary acquisition in ESP: Perspectives, strategies and resources. *Studii Și Cercetări Filologice. Seria Limbi Străine Aplicate*, 17, 171–180. Retrieved January 10, 2025, from <https://scf-lsa.info/wp-content/uploads/2019/02/162-171-OLIVIA-CHIROBOCEA.pdf>
- [13] Coxhead, A. (2012). Vocabulary and ESP. In *The Handbook of English for Specific Purposes* (pp. 115–132). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118339855.ch6>
- [14] Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications Ltd.

- [15] Dewi, S. U., & Rahma, S. (2022). Effect of using Quizlet toward students' vocabulary mastery. *Lingua-LiterA: Journal of English Language Teaching Learning and Literature*, 5(1), 33–41. <https://doi.org/10.55933/Ing.v5i1.301>
- [16] Dine, I. R. G. (2018). The use of Quizlet as an online tool to improve vocabulary mastery. *English Language Teaching Journal*, 7(5), 333–337. Retrieved February 18, 2025, from <https://journal.student.uny.ac.id/elt/article/view/13298>
- [17] Dizon, G. (2016). Quizlet in the EFL classroom: Enhancing academic vocabulary acquisition of Japanese university students. *Teaching English with Technology*, 16(2), 40–56. Retrieved March 20, 2025, from <https://tewtjournal.org/download/4-quizlet-in-the-efl-classroom-enhancing-academic-vocabulary-acquisition-of-japanese-university-students-by-gilbert-dizon/>
- [18] Flick, U. (2023). *Introduction to Qualitative Research* (7th ed.). Sage Publications.
- [19] Fox, A. (2020, March 30). *Supporting every teacher: Quizlet for teaching vocabulary*. World of Better Learning. Retrieved February 19, 2025, from <https://www.cambridge.org/elt/blog/2020/03/30/quizlet/>
- [20] Hamida, R., Prawati, A., & Hadriana, H. (2023). The effect of using Quizlet application in improving students' vocabulary mastery. *Edukatif*, 5(6), 2365–2371. <https://doi.org/10.31004/edukatif.v5i6.5681>
- [21] Hardianti, H. (2024). Gamification in EFL: Exploring the use of gamification strategies to enhance student motivation and engagement. *Intelektium*, 5(1), 8–15. <https://doi.org/10.37010/int.v5i1.1531>
- [22] Ho, D. V., Pham, T. S. T., & Thanh, N. (2020). Using Quizlet to enhance vocabulary acquisition of non-English major freshmen. *Proceedings of the 8th OpenTESOL International Conference*. Retrieved January 10, 2025, from <http://repository.vlu.edu.vn:443/bitstreams/e3f86a91-0866-4003-bbb2-b2961a8f5725/download>
- [23] Ho, T. T. H., & Kawaguchi, S. (2021). The effectiveness of Quizlet in improving EFL learners' receptive vocabulary acquisition. *Asiatic: IJUM Journal of English Language and Literature*, 15(1), 115–159. <https://doi.org/10.31436/asiatic.v15i1.2314>
- [24] Hrdličková, Z. (2020). Business and law vocabulary acquisition in the ESP classroom. *Advanced Education*, 7(14), 56–65. <https://doi.org/10.20535/2410-8286.197507>
- [25] Ika, B. (2024). Improving vocabulary acquisition: Technology-integrated, gamified, and contextualized approaches in EFL learning. *Journal of Educational Studies*, 2(2). Retrieved January 13, 2025, from <https://www.ejournal.baleliterasi.org/index.php/JEs/article/view/1144>
- [26] Jiménez-Sánchez, M., & Gargallo-Camarillas, N. (2020). Gamification and students' motivation: Using Quizizz in the English as a Foreign Language (EFL) Classroom. *Acta Marisiensis. Philologia*, 2(1), 1–13. <https://doi.org/10.2478/amph-2022-0035>
- [27] Kraus, H., Zhu, Y., & Deng, G. (2020). Gamification in large EFL classes: A preliminary investigation. *Electronic Journal of Foreign Language Teaching*, 17(2), 381–391. <https://doi.org/10.56040/hrky1732>
- [28] Liu, G., Fathi, J., & Rahimi, M. (2024). Using digital gamification to improve language achievement, foreign language enjoyment, and ideal L2 self: A case of English as a foreign language learners. *Journal of Computer Assisted Learning*, 40(4), 1347–1364. <https://doi.org/10.1111/jcal.12954>
- [29] Lubis, A. H., Johan, S. A., & Alessandro, R. V. (2022). Quizlet as an electronic flashcard to assist foreign language vocabulary learning. *Proceedings of the Sixth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2022)*, 71–76. https://doi.org/10.2991/978-2-494069-91-6_12
- [30] Lukov, T. (2022). Using Quizlet for academic vocabulary acquisition. *TESOL Working Paper Series*, 34–52. Research Publications. Retrieved February 29, 2025, from https://www.hpu.edu/research-publications/tesol-working-papers/fall-2022/3-quizlet_lukov.pdf
- [31] Luu, Q. K., & Doan, T. N. N. (2021). Use of Quizlet outside class in enhancing ESFL learners' vocabulary learning: A case study. *International Journal on E-Learning Practices (IJELP)*, 4(4), 22–32. <https://doi.org/10.51200/ijelp.v4i.3420>
- [32] Merriam, S. B. (2009). *Qualitative research*. Jossey-Bass.
- [33] Mohammadi, M., Valizadeh, M., Jalal, P. Z., & Xodabande, I. (2024). University students' academic vocabulary development through mobile-assisted learning: Exploring the impacts on receptive and productive knowledge. *Heliyon*, 10(7), e28103–e28103. <https://doi.org/10.1016/j.heliyon.2024.e28103>
- [34] Montaner-Villalba, S. (2019). The use of Quizlet to enhance vocabulary in the English language classroom. In *EUROCALL 2019* (pp. 304–309). Research-publishing.net. <https://doi.org/10.14705/rpnet.2019.38.1027>
- [35] Mykytko, I. (2023). The use of Quizlet to enhance L2 vocabulary acquisition. *Encuentro*, 31(2023), 56–69. Retrieved February 20, 2025, from <http://hdl.handle.net/10017/55791>
- [36] Nguyen, T. H. H., & Nguyen, T. D. (2021). Using Quizlet in generating learners' autonomy in learning English vocabulary. *TNU Journal of Science and Technology*, 226(03), 34–42. <https://doi.org/10.34238/tnu-jst.4100>
- [37] Nguyen, T. T., Nguyen, D. T., Nguyen, D. L. Q. K., Mai, H. H., & Le, T. T. X. (2021). Quizlet as a tool for enhancing autonomous learning of English vocabulary. *AsiaCALL Online Journal*, 13(1), 150–165. <https://doi.org/10.54855/acoj221319>
- [38] Nhac, T. H., & Ngo, H. L. C. (2023). Students' views on Quizlet as a learning medium to enhance legal English vocabulary acquisition. *VNU Journal of Foreign Studies*, 39(3), 85–101. <https://doi.org/10.63023/2525-2445/jfs.ulis.5079>
- [39] Nilubol, K. (2023). The feasibility of an innovative gamified flipped classroom application for university students in EFL context: An account of autonomous learning. *English Language Teaching*, 16(8), 24–24. <https://doi.org/10.5539/elt.v16n8p24>
- [40] Nilubol, K., & Sithitikul, P. (2025). Exploring the transformative effects of gamified learning on writing and metacognition in an EFL university context: An account of blended learning landscape. *LEARN Journal: Language Education and Acquisition Research Network*, 18(1), 513–551. <https://doi.org/10.70730/bugm2855>
- [41] Ningsih, R. P., Nasrullah, N., & Rosalina, E. (2024). The impact of Quizlet on the transferability of vocabulary in English as a foreign language learner (EFL): Student perspectives. *E3S Web of Conferences*, 483. <https://doi.org/10.1051/e3sconf/202448303013>
- [42] Normurodovna, M. A. (2025). The importance of teaching English for specific purposes. *International Conference on Scientific Research in Natural and Social Sciences*, 148–156. Retrieved January 10, 2025, from <http://econfseries.com/index.php/1/article/view/572>

- [43] Nur, S., & Putri, N. E. (2023). The Effectiveness of using Quizlet as digital learning to improve vocabulary mastery of junior high school students. *INSANIA: Jurnal Pemikiran Alternatif Kependidikan/Insania*, 28(1a), 51–63. <https://doi.org/10.24090/insania.v28i1a.8638>
- [44] Özdemir, O., & Seçkin, H. (2024). Quantifying cognitive and affective impacts of Quizlet on learning outcomes: a systematic review and comprehensive meta-analysis. *Frontiers in Psychology*, 15(2024). <https://doi.org/10.3389/fpsyg.2024.1349835>
- [45] Panigrahi, S. K. (2014). Business vocabulary acquisition: Assessing the standard of responses of vernacular and English medium students of engineering education. *Journal of Teaching English for Specific and Academic Purposes*, 2(3), 493–500. Retrieved March 20, 2025, from <https://espeap.junis.ni.ac.rs/index.php/espeap/article/view/176>
- [46] Panmei, B., & Waluyo, B. (2022). The pedagogical use of gamification in English vocabulary training and learning in higher education. *Education Sciences*, 13(1), 1–22. <https://doi.org/10.3390/educsci13010024>
- [47] Phuong, H. T. T. (2020). Gamified learning: Are Vietnamese EFL learners ready yet? *International Journal of Emerging Technologies in Learning (IJET)*, 15(24), 242–251. <https://doi.org/10.3991/ijet.v15i24.16667>
- [48] Pingmuang, P., & Koraneekij, P. (2022). Mobile-assisted language learning using task-based approach and gamification for enhancing writing skills in EFL students. *Electronic Journal of E-Learning*, 20(5), 623–638. <https://doi.org/10.34190/ejel.20.5.2339>
- [49] Platzter, H. (2020). The role of Quizlet in vocabulary acquisition. *Electronic Journal of Foreign Language Teaching*, 17(2), 421–438. <https://doi.org/10.56040/hspt1735>
- [50] Rahmani, E. F. (2020). The benefits of gamification in the English learning context. *IJEE (Indonesian Journal of English Education)*, 7(1), 32–47. <https://doi.org/10.15408/ijee.v7i1.17054>
- [51] Redjeki, I. S., & Muhajir, R. (2021). Gamification in EFL classroom to support teaching and learning in 21st century. *JEES (Journal of English Educators Society)*, 6(1), 68–78. <https://doi.org/10.21070/jees.v6i1.882>
- [52] Rofiah, N. L., & Waluyo, B. (2024). Effects of gamified grammar and vocabulary learning in an English course on EFL students in Thailand. *Teaching English with Technology*, 2024(2). <https://doi.org/10.56297/vaca6841/lrdx3699/djil1101>
- [53] Sadeghi, K., Sağlık, E., Mede, E., Samur, Y., & Comert, Z. (2022). The effects of implementing gamified instruction on vocabulary gain and motivation among language learners. *Heliyon*, 8(11), e11811. <https://doi.org/10.1016/j.heliyon.2022.e11811>
- [54] Sangtuptim, N. (2018). *Using Quizlet in vocabulary learning: Experience from Thai high school students* [Master Thesis]. Retrieved March 20, 2025, from http://ethesisarchive.library.tu.ac.th/thesis/2018/TU_2018_6021042210_10713_10581.pdf
- [55] Sanosi, A. B. (2018). The effect of Quizlet on vocabulary acquisition. *Asian Journal of Education and E-Learning*, 6(4). <https://doi.org/10.24203/ajeel.v6i4.5446>
- [56] Setiawan, M. R., & Wiedarti, P. (2020). The effectiveness of Quizlet application towards students' motivation in learning vocabulary. *Studies in English Language and Education*, 7(1), 83–95. <https://doi.org/10.24815/siele.v7i1.15359>
- [57] Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP Classroom. *RELC Journal*, 45(1), 51–65. <https://doi.org/10.1177/0033688214522641>
- [58] Tayeh, Q., Krishan, T. M., & Malkawi, N. (2024). The effect of using gamification to improve EFL students' academic performance. *Journal of Ecocumanism*, 3(7), 45–54. <https://doi.org/10.62754/joe.v3i7.4173>
- [59] Todorova, N. (2023). Effect of Quizlet use on Ukrainian service members' English vocabulary acquisition. *Advanced Education*, 11(23), 29–41. <https://doi.org/10.20535/2410-8286.281048>
- [60] Tran, D. T., & Le, T. H. (2023). The impact of Quizlet on vocabulary improvement: A case study in Binh Duong province secondary schools. *World Journal of English Language*, 13(6), 235–235. <https://doi.org/10.5430/wjel.v13n6p235>
- [61] Tran, M. D., & Nguyen, T. M. L. (2022). Improving students' vocabulary retention through using Quizlet: An action research at high school. *TNU Journal of Science and Technology*, 227(04), 170–177. <https://doi.org/10.34238/tnu-jst.5647>
- [62] Tran, T. Q., & Nguyen, L. T. (2023). Insights into Vietnamese bilingual students' Quizlet-based English vocabulary learning. *Journal of Language Teaching and Research*, 14(3), 682–688. <https://doi.org/10.17507/jltr.1403.15>
- [63] Trinh, P. H., & Ho, T. B. M. (2023). Quizlet's impacts on EMI students' vocabulary mastery and lesson comprehension at Hanoi University of Industry: A student perspective. *Journal of Science and Technology*, 59(6B). <https://doi.org/10.57001/huih5804.2023.241>
- [64] Václavík, L. (2020). The role of Quizlet in learning business vocabulary. *CASALC Review*, 10(1), 69–88. <https://doi.org/10.5817/casalc2020-1-6>
- [65] Vogel, M. (2022, May 3). *Quizlet for learning vocabulary*. Active Learning at King's. Retrieved January 10, 2025, from <https://blogs.kcl.ac.uk/activelearning/2022/05/03/quizlet-for-learning-vocabulary/>
- [66] Waluyo, B., & Bucol, J. L. (2021). Does gamification work for low proficiency students? Gamified vocabulary at home and learning outcomes in class using Quizlet. *Computer-Assisted Language Learning Electronic Journal*, 22(1), 164–185. Retrieved March 20, 2025, from <https://callej.org/index.php/journal/article/view/326>
- [67] Wright, B. A. (2016). Transforming vocabulary learning with Quizlet. *JALT2016 • Transformation in Language Education*. Retrieved January 10, 2025, from <https://jalt-publications.org/node/4/articles/6075-transforming-vocabulary-learning-quizlet>
- [68] Wulandari, M. F., & Irfani, B. (2023). Enhancing business management students' English vocabulary through Quizlet: Exploring classroom action research. *Journal of English Teaching*, 9(3), 350–360. <https://doi.org/10.33541/jet.v9i3.5157>
- [69] Zakian, M., Xodabande, I., Valizadeh, M., & Yousefvand, M. (2022). Out-of-the-classroom learning of English vocabulary by EFL learners: Investigating the effectiveness of mobile assisted learning with digital flashcards. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1). <https://doi.org/10.1186/s40862-022-00143-8>
- [70] Zhang, S., & Hasim, Z. (2023). Gamification in EFL/ESL instruction: A systematic review of empirical research. *Frontiers in Psychology*, 13, 1–12. <https://doi.org/10.3389/fpsyg.2022.1030790>
- [71] Zhang, Z., & Crawford, J. (2023). EFL learners' motivation in a gamified formative assessment: The case of Quizizz. *Education and Information Technologies*, 29, 6217–6239. <https://doi.org/10.1007/s10639-023-12034-7>
- [72] Zhou, S. (2024). Gamifying language education: the impact of digital game-based learning on Chinese EFL learners. *Humanities and Social Sciences Communications*, 11(1). <https://doi.org/10.1057/s41599-024-04073-3>

Thuan D. Pham is a senior English lecturer at Department of Foreign Languages and Information Technology, Hoa Lu University, Viet Nam. He received his Ph.D. in English language teaching methodology from University of Languages and International Studies, Vietnam National University, Hanoi, in 2022. He has more than ten years of experience teaching English at the tertiary level. He is interested in TESOL, learner autonomy, professional development, CALL, and MALL (Email: pdthuan@hluv.edu.vn; ORCID iD: <https://orcid.org/0000-0001-8535-4411>).

Tam T. Pham is a senior English lecturer at Department of Foreign languages, Academy of Finance, Ha Noi, Vietnam with more than fifteen-year experience at the tertiary level. She received her Ph.D. degree in English linguistics from University of Languages and International Studies, Vietnam National University, Hanoi in 2023. She has a passion for teaching students English for their future career. She is particularly interested in professional development, CALL, cultural studies and interdisciplinary research (culture and finance) (Email: phamthitam@hvtc.edu.vn).

Ecocentrism, Irony, and the Wounds of Nature in the Poem “Aku Ini”: Ecocriticism in the Context of North Maluku

Wildan

Study Program of Indonesian Literature, Universitas Khairun, Ternate, Indonesia

Ramis Rauf*

Study Program of Indonesian Literature, Universitas Khairun, Ternate, Indonesia

Ratna

Study Program of Indonesian Literature, Universitas Khairun, Ternate, Indonesia

Afriani Ulya

Study Program of Indonesian Literature, Universitas Halu Oleo, Kendari, Indonesia

Vidhia Agmareina Hirto

Study Program of Travel and Tourism Business, Universitas Khairun, Ternate, Indonesia

Sulmi Magfirah

Study Program of English Literature, Universitas Khairun, Ternate, Indonesia

Fitriningsih Pratiwi Mahmud

Study Program of English Literature, Universitas Khairun, Ternate, Indonesia

Dhini Yustia Widhyah Saputri

Study Program of Indonesian Literature, Universitas Khairun, Ternate, Indonesia

Ulfah Dwi Januarty

Study Program of Indonesian Literature, Universitas Khairun, Ternate, Indonesia

Ahmad Rifani Talaohu

The University of Newcastle, Australia;

Study Program of English Literature, Universitas Khairun, Ternate, Indonesia

Abstract—The poem "Aku Ini" is from the anthology *Karang Menghimpun Bayi Kerapu* by Ibrahim Gibra. This poem conveys ecological irony: humans are praised for their clever symbolic "embroidering" of nature, yet they fail to sustain the environment. This study uses an ecocritical approach in a descriptive-qualitative design with textual close reading of diction, metaphors, stanza structures, and natural imagery. The main data, in the form of poetic texts, is contextual data that includes depictions of ecological degradation in North Maluku (deforestation, coastal pollution, plastic waste, mining pressure) as a background for the reading. The analysis maps the imagery of dew, sea, river, and socio-political realm (democracy) into Garrard's ecocritical theme clusters—pollution, pastoral, wilderness, animals or non-human agents, and the earth—to examine the construction of human-nature relations and traces of anthropocentric ideology. In the research, the findings show that there are three main elements, namely (1) the poem constructs a landscape of ecological wounds through the metaphor of a damaged natural body, (2) the repetitive irony of "cleverness" versus failure to act becomes a critique of ecological indifference and the decline of public ethics, and (3) the lyrical voice opens a shift towards eco-centric sensibility and a reflective call to ethical responsibility across creatures. In an eco-social functional way, this poem serves as a cultural warning and a pedagogical medium to foster ecological awareness in Indonesian literary studies, especially those rooted in environmental experiences in North Maluku.

Index Terms—ecocritical, irony, ecology, ecocentrism, North Maluku

I. INTRODUCTION

Systematic environmental damage has occurred since the Industrial Revolution era in the 18th century (Jambeck et al., 2015), but has become more widespread and become a central issue in global discourse in the mid-20th century (Carson, 1962; Shiva, 2005a; World Bank, 2021; Guha, 2000; UN, 2022), not only in science and politics, but also in literature (Gore, 2006; Plumwood, 2002; Merchant, 1980). Literature has the power to reflect and criticise ecological issues through (poetic) language, as stated by Glotfelty and Fromm (1996) and Buell (2005), that literature can act as a medium, a mouthpiece that voices the ecological crisis and encourages strong environmental awareness through aesthetic representation. In this context, Ibrahim Gibra's poem "Aku Ini" serves as a compelling example of how literary works can be a space for critiquing the environmental crisis.

In Indonesia, environmental damage is inextricably linked to various massive natural resource extraction practices, including mining and waste management (Rauf et al., 2024b). In North Maluku, for example, the expansion of the nickel and gold mining industry has led to environmental degradation and even destruction, including deforestation, river pollution, and coastal habitat destruction (Jambeck et al., 2015; United Nations Environment Programme, 2022; World Bank, 2021). Consequently, these mining activities not only damage local ecosystems but also give rise to social conflicts that threaten the sustainability of indigenous (local) communities that are highly dependent on nature (Li, 2014; Rangkuti, 2009). This reality provides a relevant contextual background for interpreting the poem "Aku Ini" as a critique of humans' exploitative relationship with nature.

Aesthetically, this poem depicts "Me," an intelligent individual capable of symbolically interacting with natural elements, such as forests, seas, and rivers (Taha et al., 2024). However, the figure of "I" actually demonstrates a failure to carry out concrete responsibilities towards the environment. The discrepancy between rhetoric and real action is the main criticism of the poem "Aku Ini," offering a space for reflection and prompting readers to realise the paradox of humanity in an ecological context: the imbalance between intellectual claims and ethical responsibility towards nature. More exclusively, the poem "Aku Ini" presents a representation of nature through powerful symbolism: deforested forests, seas strewn with plastic waste, and rivers clogged with waste—a landscape of ecological wounds conveyed aesthetically and politically (Subuh et al., 2024). In this context, nature is not merely a backdrop but also a suffering subject, voicing its grievances through humanity's failure to maintain balance in life (Rauf et al., 2025). This symbolism conveys the profound impression that environmental damage reflects humanity's moral and spiritual decay. Forests, seas, and rivers become metaphors for the torn body of the world, which harbours a narrative of profound pain caused by human arrogance (Rasiah et al., 2024).

The critique of indifference is a dominant theme in this poem. When "I" claims to be intelligent but fails to plant trees, or is eloquent in speaking but unable to save rivers, what results is a confirmation of the distortion of reality, the moral nakedness of modern humans—skilled at stringing words together but dull in concrete action regarding environmental concerns. This situation is sufficient to illustrate satire of contemporary society, which is more preoccupied with rhetoric than with concrete action to save or heal the wounded (damaged) earth. This indifference appears as a form of human alienation from nature, a previously harmonious relationship, but has now turned into an exploitative relationship.

The poem "Aku Ini" exudes eco-centric awareness, a view that places nature as an entity with intrinsic value, not merely an object of human exploitation. Each stanza implicitly calls for humanity to review nature as an integral part of its existence, rather than as an enemy or a tool. In "Aku Ini," ecological awareness is not presented declaratively but conveyed through irony, regret, and a poignant, reflective awareness. This ecocentrism challenges the dominant anthropocentric view that places humans at the centre of all events, while calling for reconciliation between humans and the earth as a shared ethical community (Rauf et al., 2024b).

In light of this awareness, the urgency of examining the poem "Aku Ini" becomes even more apparent. This research is crucial because the poem not only addresses nature aesthetically but also voices ecological concerns through powerful social critique. In the context of growing global awareness of the impacts of climate change, deforestation, marine pollution, and other forms of environmental degradation, literary works that address ecological themes can serve as instruments for fostering collective awareness. This awareness of ecological urgency aligns with Buell's (2001) view, which emphasises that literary works intervene in how society views and treats nature. Furthermore, an ecocritical approach to poetry analysis is important because it provides an interdisciplinary perspective that connects literature, ethics, and ecology. In the Anthropocene era, when human activity has become a dominant geological force transforming the planet's ecosystems, literary works such as the poem "Aku Ini" can serve as a means to contemplate the moral crisis of modern humanity in the face of ecological destruction (Heise, 2006). Therefore, this study is highly relevant to literary literacy, supporting the discourse on sustainable development and interdisciplinary ecological responsibility.

The novelty of this research lies, first, in the application of a contextual ecocritical approach to contemporary Indonesian poetry that addresses environmental damage in North Maluku. Second, this research integrates poetic symbolism, such as dew, the sea, and rivers, with the real ecological crisis caused by extractive activities, and interprets irony as an aesthetic expression of human ethical failure to preserve nature. Third, it positions North Maluku as an ecological-cultural space that integrates environmental issues and local politics into the literary repertoire, thereby broadening the scope of ecocriticism to encompass the social and geopolitical realms of eastern Indonesia. By considering several ecological explanations, this study is expected to contribute to the development of ecocritical studies and emphasise the important role of literary works, particularly poetry, as agents of social and ecological change. This

approach emphasises how literature is not merely a reflection of reality but also capable of provoking awareness and encouraging transformative action in response to the global environmental crisis.

II. LITERATURE REVIEW

The theory used in this research is ecocriticism. Ecocriticism is a branch of literary studies that has developed since the 1990s and focuses on the relationship between literary texts and the natural environment, both physically and ideologically. According to Garrard (2004), ecocriticism examines how nature is represented in literary works and how the relationship between humans and the environment is constructed, critiqued, and interpreted within the text. Beyond simply reading natural settings descriptively, ecocriticism explores the ethical, political, and symbolic dimensions of the human-nature relationship deliberately presented in literary works. Ecocriticism also assumes that the ecological crisis is not merely a scientific or technical issue but also a matter of culture and imagination. Therefore, literature is seen as having great potential to shape public ecological awareness through reflective, engaging narratives. In this context, ecocritical theory not only examines environmental representations but also uncovers the anthropocentric ideology inherent in the text, affirming eco-centric values that give space to nature as a living entity equal to humans (Buell, 2005; Glotfelty & Fromm, 1996).

In this study, Garrard's ecocritical tools are operationalised to map the poem "Aku Ini's" natural images into key thematic clusters, including pollution, wilderness, and the earth. These themes should inspire the audience to see the connection between local ecological issues and broader global environmental concerns. The next step is an ideological reading, examining the extent to which the lyrical voice reproduces or challenges anthropocentrism, opening up the possibility of an eco-centric reading that gives symbolic agency to non-human elements (dew, sea, river). This approach facilitates the relationship between textual representation and ecological experiences in North Maluku, including Central Halmahera, East Halmahera, North Halmahera, and South Halmahera—particularly ecological pressures resulting from extractive activities and coastal landscape changes—as part of a cultural discourse that can shape ecological awareness. For example, in 2023, the local indigenous people living near the Sagea River faced a change in the freshwater. The mining exploration in the Weda district causes it. Thus, the acceleration of ecocritical theory here does not stop at describing natural settings, but also works as a tool of ethical and political criticism that negotiates local environmental values within global ecocritical conversations (Buell, 2005; Garrard, 2004; Glotfelty & Fromm, 1996).

Operationally, each of Garrard's mapped themes is used as a reading lens: (1) pollution is characterised by the diction of damage, waste, and wounds of nature; (2) pastoralism highlights the contrast or longing for relatively intact natural spaces compared to modernity or industrial exploitation; (3) wilderness evokes the image of wild areas or spaces beyond human control; (4) animals present non-human figures as witnesses, victims, or ecological partners; and (5) the earth opens up broader ecological horizons—collective, planetary, or cosmic. This mapping helps to trace the layers of ecological irony in the poem, namely when poetic language simultaneously celebrates nature and hints at its degradation, and allows for a reading of the shift from an anthropocentric attitude towards an eco-centric sensibility in the lyrical voice (Garrard, 2004). Modern ecocritical studies often draw on the work of Greg Garrard, who maps key themes such as pollution, wilderness, animals, and the earth, and how these themes operate within literary texts and cultural discourse (Garrard, 2004). Garrard's categorical framework provides a useful reading tool for analysing the symbols of nature (dew, sea, river) in the poem "Aku Ini." However, unlike Garrard's focus on Western traditions, this study applies these tools to contemporary Indonesian lyric poetry and anchors it in ecological damage in North Maluku.

The initial foundations of ecocriticism were laid in the anthology "The Ecocriticism Reader," edited by Glotfelty and Fromm (1996). In this context, ecocriticism is the study of the relationship between literature and the physical environment, and it asserts that literary texts can serve as vehicles for ecological awareness. This study builds on this basic idea by demonstrating how Ibrahim Gibra's poetry—born within the socio-environmental landscape of eastern Indonesia—functions as contextual ecological critique, not merely aesthetic expression. Two works by Lawrence Buell enrich the theoretical foundation. In "The Future of Environmental Criticism," Buell (2005) calls for a more global, cross-cultural ecocriticism and emphasises the role of the literary imagination in shaping public responses to environmental crises. Elsewhere, in *Writing for an Endangered World* (2001), Buell introduces the notion of toxic discourse—the way literature mobilises empathy through depictions of threatened places. This study actualises both of Buell's ideas by reading the line "sea wallowing in plastic" as a toxic image, while also placing regional poetry (North Maluku) within the global ecological conversation envisioned by Buell.

Heise (2006) emphasises, through the concept of eco-cosmopolitanism, the need to bridge the "sense of place" (local attachment) with the "sense of planet" (global awareness). This framework is highly relevant for current research linking local symbols—dew, sea, rivers—in Ibrahim's (2019) poetry with global issues such as plastic pollution and climate change. In line with this view, Naess (1973) distinguishes shallow anthropocentric ecology from deep ecology, which recognises the intrinsic value of all creatures. Naess's eco-centric notion helps interpret the failure of the "I" who "cannot pick dew" or "neglects to plant trees" as an ecological ethical crisis in the poem.

Shiva (2005a) emphasises the dimension of ecological justice in *Earth Democracy*, showing how the expansion of extractive and corporate capital marginalises local communities and damages ecosystems. This research translates Shiva's critique into a poetic reading. The poem's stanza "democracy" is read as an ironic commentary on the promise of political representation that collapses when mining, deforestation, and pollution injure community living spaces. The perspective

of resource political ecology is deepened through Li's (2014) ethnography, which documents how the penetration of extractive capital changes indigenous peoples' relationships with the land. While Li's work draws on socio-economic field data, this research focuses on absorbing that sensibility to read the echoes of ecological conflict in the symbolic imagery of the poem "Aku Ini".

The empirical aspect of the environmental crisis—particularly marine plastic pollution—is underscored by global research by Jambeck et al. (2015), which estimates that millions of tons of plastic waste leak from land into the ocean each year, particularly from coastal areas with weak waste management. The United Nations Environment Programme Report, *From Pollution to Solution* (2021), confirms these findings, noting that plastic dominates marine debris and has widespread impacts on ecosystems, the economy, and health. This study did not conduct biophysical measurements but utilised the data to argue that the image of a "sea wallowing in plastic" in poetry is not aesthetic hyperbole but rather the poetic resonance of a real ecological crisis, including on the coast of North Maluku.

Three additional references help position this study within a broader context. Guha (2000) traces the history of the global environmental movement with an emphasis on the experiences of the Global South. His historical perspective helps situate the ecological case in North Maluku within a long line of resource conflicts. Gore (2006) popularised the moral urgency of the climate crisis with an audience. This urgency is rhetorically adopted in this study, where he emphasises the importance of reading poetry as an ecological alarm. Merchant (1980) in *The Death of Nature* critiques modern mechanistic rationality that objectifies nature. This study demonstrates the echo of this objectification in lyrical form—nature as a body wounded by human arrogance—through the voice of the "I" in poetry.

Overall, the theoretical literature on ecocriticism (Buell, 2005; Cislighi & Heise, 2020; Garrard, 2004; Glotfelty & Fromm, 1996) provides a strong conceptual framework, while studies of ecological justice and political ecology (Li, 2014; Shiva, 2005b) highlight power structures in resource exploitation. Environmental science studies (Rauf et al., 2024b) supply empirical data on the coastal and plastic crises. However, studies rarely bring these three together in a reading of Indonesian poetry anchored in a specific ecological case. It is within this open space that this research takes on a role, namely integrating poetic symbols, the context of environmental damage in North Maluku (mining, plastic pollution, river and coastal degradation), and an ethical-ecological critique to broaden the horizons of ecocriticism in Indonesia.

III. METHODOLOGY

This study employed a descriptive-qualitative method with a textual approach. The primary data source was the poem "Aku Ini" by Ibrahim (2019). The analysis focused on the poem's diction, metaphor, and structure, as they relate to ecological themes. The interpretation process was conducted by referring to basic ecocritical concepts through careful, meticulous, and repeated readings of the poem "Aku Ini," particularly those related to environmental degradation, the paradoxes of modernity, and critiques of anthropocentric culture. An alternative approach was a descriptive-qualitative approach, which is considered capable of exploring the deeper meanings of the symbols and natural imagery contained in the poem "Aku Ini." This method allowed the analysis to go beyond textual aspects, such as diction or structure, to broadly interpret the connotative dimensions of meaning through an ecocritical theoretical framework. Close reading techniques were used to identify patterns of natural representation and ecological critique, which were then linked to the socio-cultural context and environmental degradation of North Maluku. This process enabled researchers to examine how metaphors, ironies, and paradoxes in Ibrahim's poetry reflect the complex relationships among humans, modernity, and nature.

IV. DISCUSSION

The poem "Aku Ini" possesses complex and reflective layers of meaning. With a lyrical, contemplative style, this poem depicts the figure of the "I," praised for its symbolic ability to interact with nature. However, this style also demonstrates its practical failure to care for and protect the environment. In each stanza, this poem emphasises the irony that social recognition of the "I" as intelligence and the ecological reality demonstrate neglect, or even powerlessness, in resisting. By positioning literature (poetry) as critique, the poem "Aku Ini" offers a sharp critique of humanity's relationship with nature, especially in the face of technological advances in the modern, global era, which are rife with rhetoric but lack concrete action. Through dense diction and symbolism, the poet weaves images of deforested forests, polluted seas, and rivers choked with waste, all presented as landscapes of ecological wounds rooted in human negligence.

A. Representation of Nature and Symbolism

The poem "Aku Ini" consists of four stanzas, each presenting a natural element as the main symbol: dew, the sea, the river, and democracy as a socio-political space that should support the sustainability of life. These four keywords represent aspects of nature and social life that are being or have been harmed, including forests losing their trees, seas filled with plastic, and rivers trapping garbage due to a lack of genuine human concern. Referring to the title, "I" is an individual who asserts himself over the existence of attitudes trapped in praise or arrogance that have lost face. In each stanza, "I" claims to be good at speaking, parsing words and using beautiful diction to invite readers to communicate—"embroidering nature," "calling the sea," "standing in front of the river," "arranging discourse"—but all of this fails to be realised in real

action. These symbols indicate a paradoxical content or conflict between the symbolic and practical, specifically the ability to speak without concern for nature and others (colleagues, individuals, society, and/or communities).

Dew symbolises the hope for sustainability and greenery in a beautiful yet fragile life. In the verse, “People say, I am skilled at embroidering nature, but I cannot even pick a flower from a single dewdrop,” the failure of “myself” to pick a flower from a single dewdrop indicates an inability to maintain nature’s delicacy and balance. In the ecological context of the North Maluku community, dew can be interpreted as a symbol of local wisdom and the ecological spirituality of indigenous peoples, as well as the balance of mountain ecosystems and tropical forests, such as those found in Halmahera. This ecological space has long supported the region’s biodiversity and microclimate. The failure of “myself” to pick a flower from the dewdrops reflects humanity’s inability to sustain natural ecosystems, which are fragile due to illegal logging, nickel mining, and forest conversion for extractive industries. Dew, which should be part of the natural water cycle and help cool a forest microclimate, has now become a symbol of an ecosystem nearly lost due to the destruction of vegetation cover, disrupting the eco-hydrological balance. In this context, the dew in the poem “Aku Ini” speaks not only to an aesthetic exploration of words but also to a warning about the fragility of North Maluku’s natural environment, which is constantly threatened by a lack of ecological awareness. It is a form of ecological disconnection or alienation from nature, meaning a condition in which humans no longer feel connected to nature, either emotionally, spiritually, or practically.

In the second stanza, the sea, which is generally a symbol of longing and inner adventure, is presented as an ocean “walled in plastic.” In this stanza, anxiety is expressed. In the verse “but even on a line of the coast I cannot roll my longing/in a sea walled in plastic,” the once romantic sea now marks an ecological crisis caused by plastic waste pollution. In the concrete context of coastal ecology, such as on the islands of Ternate and Tidore and along the coast of Halmahera in North Maluku, the sea “walled in plastic” reflects an increasingly real reality. It means that the once clean coastline is now filled with household waste, single-use plastic packaging, and even ship waste. The sea, which has historically been a source of livelihood for fishermen and a spiritual space for coastal communities, has now lost its sanctity due to consumption and poor waste management. The phrase “I cannot roll my longing” serves as a metaphor for human longing for a clean sea, which is currently difficult to achieve because environmental pollution has altered the ecological landscape and the aesthetics of the coast. At various points along the coast of North Maluku, piles of plastic waste carried by currents and/or flowing from settlements in the highlands have accumulated, a symptom of the flow of waste that threatens coral reef ecosystems, marine life, and even the livelihoods of fishing communities. In the poem “Aku Ini,” the sea is no longer a calm inner refuge, but rather a mirror of a terrifying ecological wound born of human indifference.

In the repetition of the complaint in the third stanza, the river reveals the problem of the garbage-labyrinth, depicted as a source of life that is now dirty and no longer suitable for sustaining life. The metaphorical phrase “I neglected to plant trees, then the thick currents trap the garbage” illustrates human negligence towards environmental concerns, an attitude that transforms the river from a source of life to a garbage dump, rendering it an unhealthy place to sustain life. In an ecological context, especially in the inland areas of South Halmahera, Obi, and the surrounding areas, the river that was once clear and a primary source of clean water is now polluted by mining activities, illegal logging, and domestic waste disposal. The phrase “I neglected to plant trees” refers to the loss of vegetation in the upstream area, which leads to severe erosion and sedimentation. At the same time, “the thick currents trap the garbage” concretely depicts the reality of small rivers that are now full of mud, household waste, and inorganic waste. In this stanza, the river is no longer a source of life. However, it has become a waste channel that threatens public health and the sustainability of the freshwater ecosystem. Through the metaphor of the poem “Aku Ini,” the river serves as a signifier, namely a symbol of ecological damage caused by collective human negligence, including societal neglect, corporate negligence, and weak state oversight of activities that harm river basins.

Democracy, as a socio-political symbol, is presented as meaningless. In the final stanza, “even in a single piece of democracy, I cannot weave joy in words that have lost their face,” words lose their value because they become mere rhetoric (empty), statements that are not followed by actions that side with the people or nature. In the socio-political context of eastern Indonesia, including North Maluku, democracy often exists only in procedural form—elections take place, directed representatives are elected—but is not followed by real support for actions in the interests of the people, especially on ecological issues. The phrase “words that have lost their face” reflects the reality that many political promises related to environmental protection are not kept, for example, in the case of mining permits granted in customary areas and protected forests, without community consent. When local communities voice opposition to mining or deforestation, their aspirations are often ignored or silenced by law enforcement officials and civil servants. In this context, democracy becomes a symbol of a system that has lost its conscience, as it no longer accommodates society’s ecological aspirations but instead supports the exploitation of nature. The text, the poem “Aku Ini,” in this context reflects the loss of substance in political discourse: language no longer liberates and protects, but merely serves as a tool to justify policies that harm the earth.

Within the realm of natural representation and symbolism, this poem combines natural and social symbols to portray ecological damage and the human condition on the verge of a moral crisis. Through “the self,” modern humans are represented as individuals or groups who like to make promises, full of promises, but reluctant to take real action to protect the earth. Nature, which should be guarded, cared for, and protected as a form of respect, instead becomes a silent witness to the occurrence of destruction and indifference. The symbolism of the poem “Aku Ini” has inspired readers,

particularly in contemplation, to realise that love for nature cannot be expressed solely through words; it must also be demonstrated through concrete actions.

B. Criticism of Indifference

In the poem “Aku Ini,” diction such as “I cannot pick flowers,” “I cannot roll away my longing,” and “I neglect to plant trees” convey a sense of powerlessness or reluctance that leads to apathy caused by a loss of hope, an attitude fueled by the failure of great men who were expected to bring new hope to take concrete action to save the environment. Through “Aku Ini,” the poem critiques modern human culture, which lives in a global era characterised by excessive talk and rhetoric, but minimal action. This view is emphasised through the diction “democracy” and “words that have lost their face,” which reinforce the social critique that even the system that should represent the will of the people has lost its meaningful value and its primary nuances of meaning.

In the poem “Aku Ini,” each diction represents a painful ecological reality and human negligence in preserving nature. The phrase “I cannot pick flowers” signifies the damage that has occurred so that “I have lost the ability and motivation to enjoy the beauty of nature”; “I cannot roll my longing” symbolises the emotional disconnection (intimacy) between humans and nature. Meanwhile, “I neglected to plant trees” is an expression of honesty, an admission of negligence in environmental conservation actions. In this phase, these sentences reflect the character of humans who are more busy constructing verbal narratives than taking real action. The symbol of “democracy” and the critique of “words that have lost face” emphasise that ecological issues certainly do not stand alone. Ecological destruction is closely related to the decline of public ethics and the collapse of public trust in the social and political systems that supposedly guarantee the sustainability of living together in a shared space, namely, a peaceful nature.

The critique of indifference in the poem “Aku Ini” principally targets two groups: individuals and social and cultural structures that perpetuate the imbalance between discourse and action. The “I” as a signifier represents the figure of modern humanity, possessing symbolic awareness but weak and blunted in ecological responsibility. When words lose their meaning and actions fail to follow suit, the environmental crisis is further amplified. The phrase “I neglect planting trees” in this context is not merely a metaphor but a concrete depiction of the loss of human empathy for the regeneration process or for the protection of nature. This statement demonstrates that environmental damage is not solely the result of industrial activity but also stems from apathy and a weakening moral awareness of the need to preserve nature. As a form of critique, the poem “Aku Ini” serves as a stern rebuke to modern culture, which tends to worship discourse while neglecting concrete action—a form of ecological hypocrisy that exploits nature for the benefit of oligarchs or certain groups (Rauf et al., 2024b).

The poem “Aku Ini” is a sharp critique of the apathy of modern humans living in the era of globalisation, depicted through the lyrical voice of the self. Diction or phrases such as “I cannot pick flowers,” “I cannot roll up my longing,” and “I neglect to plant trees” serve as metaphors for humanity’s inability to enjoy, care for, and sustain nature. These expressions signify the loss of emotional intimacy with nature; the self not only feels disconnected from the beauty of the environment but also acknowledges its negligence in taking concrete action to preserve the ecosystem. This critique is emphasised through the symbols of “democracy” and “words that have lost their face,” words that indicate that the social and political systems have failed to fulfil their ecological responsibilities. Democracy, which should be a space for public participation for the common good, has instead lost its substance—the sense of democratic togetherness, promoted from the beginning, is merely empty rhetoric deliberately designed to trap, unaccompanied by concrete steps to preserve nature. In this context, Ibrahim’s poem, through its juxtaposition of ecological critique with socio-political critique, demonstrates that the environmental crisis is inextricably linked to the decline of public ethics and the collapse of public trust in power structures.

This attitude is evident in the “I,” whose actions symbolise modern humans who are aware of their consciousness but cannot take concrete action. The phrase “I neglect to plant trees” is not merely a metaphor, but a concrete reflection of moral failure and a loss of empathy for the natural regeneration process. Environmental damage, as depicted in this poem, is not only the result of industrial exploitation but also arises from a human attitude that worships words and discourse while rejecting ecological responsibility. Through the “I,” the poem “Aku Ini” not only rebukes individuals but also exposes the hypocrisy of modern culture. This culture exploits nature for the benefit of a small group (oligarchy) while disregarding the ecological suffering it causes.

C. Ecocentrism and Ecological Awareness

The poem “Aku Ini” radiates eco-centric values, namely, viewing nature not as an object to be exploited but as an equal subject. The emphasis on the loss and damage to nature indicates that the poet is not worshipping nature in a romantic, bombastic manner, but rather is voicing a truly critical ecological crisis. Ecological awareness is apparent through the irony and regret of “the I” in all stanzas, marked by the choice of the word “clever,” which represents modern humans who are always in discourse but minimal in real action. Eco-centric awareness in the poem “Aku Ini” is evident in the depiction of nature as an entity with life, feelings, and the right not to be harmed. In a series of stanzas, the poet depicts nature in two aesthetic orientations: the setting or resources, and nature as a body that is experiencing injury and pain due to human actions. Expressions such as “I neglected to plant trees” or “the sea has become a place to dump my failed longing” show three interrelated signs: environmental damage in physical, moral, and spiritual dimensions. The poet presents the human-nature relationship as broken—not due to ignorance, but rather to a deliberate disregard for ecological

responsibility. In the eco-centric perspective, Naess (1973) and Plumwood (2002) have positioned nature as a subject suffering due to human arrogance and a sense of superiority (Glotfelty & Fromm, 1996).

This poem demonstrates the irony of a profound gap: modern humans in a globalised world often feel capable of explaining and controlling nature, yet, emotionally, they have lost their sense of intimacy. Words like “I embroider the dew,” “calling the sea,” and “departing the river” reflect the symbolic ability to touch or speak with nature, but all of this is merely rhetoric and empty talk. This symbolism suggests that the human relationship with nature is reduced to language rather than to action, discourse, or genuine concern. This irony reinforces the impression that humans have lost their intimacy and respect for nature as fellow creatures in the web of life.

The ecological awareness in the poem “Aku Ini” is demonstrated through regret and the “I” delay in realising the impact of the wounds inflicted on nature. This regret is felt in the nuance of a quiet confession that all the words ever spoken are not enough to heal the bare forests with gaping wounds, the seas filled with garbage, or the rivers that are lifeless from clogging with waste. This awareness is a form of critical reflection on the anthropocentric paradigm, namely, the positioning of humans as the centre of the universe.

The poem “Aku Ini” represents an eco-centric view in which the main axis of life (nature) is no longer positioned as a passive object to be exploited, but as a living subject with the right to be respected. This eco-centric value is reflected in the criticism of the damage and loss of nature, which is depicted not as romantic nostalgia, but as a sign of a real, critical ecological crisis. Ecological awareness is evident in the irony and regret of the “I” in almost all stanzas, particularly in the word “clever”—a satire on modern humans who are rich in discourse but poor in action. In this context, the poet presents nature as a living entity that experiences wounds and suffering due to human greed. Diction such as “I neglected to plant trees” or “the sea became a place to dump my failed longing” reveals three layers of wounds: physical (environmental damage), moral (negligence of responsibility), and spiritual (loss of intimacy with nature). This entity aligns with the eco-centric concept of Arne Naess and Plumwood, which posits that nature is an integral part of the network of life, equal to humans. In terms of description and argumentation, the poem “Aku Ini” serves as a rebuke to anthropocentric superiority that justifies the exploitation of nature in the name of progress.

V. CONCLUSION

The poem “Aku Ini” offers sharp, reflective ecological criticism. Through an ecocritical approach, this poem demonstrates that modern humans are not only clever at talking about nature but also neglectful of protecting and preserving it. In “Aku Ini,” nature is depicted as an entity wounded by the absence of concrete human action to care for it. The poem “Aku Ini” emphasises that language and symbols are insufficient means or markers of love for nature without a commitment to and concrete action to preserve the sustainability of life. An analysis of the poem “Aku Ini” demonstrates that literature is not merely a medium for aesthetic expression but also a tool for social criticism and ecological reflection. This poem presents a wounded ecological reality through irony, metaphor, and poignant symbolism, thereby touching readers’ ethical awareness of the environmental crisis. These findings have theoretical implications for strengthening the ecocritical approach in literary studies, positioning it not merely as a complement to textual analysis but rather as a more robust theoretical framework that is relevant and adaptive in facing the challenges of our time, particularly the climate crisis and environmental damage caused by modernisation and the exploitation of natural resources in North Maluku.

This article is a concrete step toward raising environmental awareness among the younger generation in North Maluku, especially among those majoring in literature. By studying ecocriticism, they will understand how literature can be used to promote environmental advocacy, helping them become environmental agents who will educate their relatives—who live around mining areas across almost all of Halmahera Island—about the prohibition on destroying nature and the call to preserve forest ecosystems to avoid natural disasters such as floods and landslides. We can give an example, such as what happened in 2023, when the Sagea River—located below Bokimaruru Cave, North Weda District, Central Halmahera Province—was polluted by mining waste generated by several mining companies operating around the river. The water of the Sagea River, which was originally clean, clear, and fresh, turned brown. As a result, indigenous communities experienced difficulties in obtaining clean water.

REFERENCES

- [1] Buell, L. (2001). *Writing for An Endangered World: Literature, Culture, and Environment in the U.S. and Beyond*. Cambridge: Harvard University Press.
- [2] Buell, L. (2005). *The Future of Environmental Criticism: Environmental Crisis and Literary Ecology*. Oxford: Blackwell Publishing.
- [3] Carson, R. (1962). *Silent Spring*. Boston: Houghton Mifflin.
- [4] Cislighi, B., & Heise, L. (2020). Gender Norms and Social Norms: Differences, Similarities and Why They Matter in Prevention Science. *Sociology of Health & Illness*, 42(2), 407–422. <https://doi.org/10.1111/1467-9566.13008>
- [5] Garrard, G. (2004). *Ecocriticism*. London: Routledge.
- [6] Glotfelty, C., & Fromm, H. (1996). *The Ecocritism Reader: Landmarks in Literary Ecology*. Athens: Georgia University Press.
- [7] Gore, A. (2006). *An Inconvenient Truth: The Planetary Emergency of Global Warming and What We Can Do About It*. Pennsylvania: Rodale.
- [8] Guha, R. (2000). *Environmentalism: A Global History*. Londong: Longmann.

- [9] Heise, U. K. (2006). *Sense of Place and Sense of Planet: The Environmental Imagination of the Global*. Oxford: Oxford University Press.
- [10] Ibrahim, G. A. (2019). *Karang Menghimpun Bayi Kerapu* [Coral Gathers Baby Grouper]. Yogyakarta: JBS.
- [11] Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., Narayan, R., & Law, K. L. (2015). Plastic Waste Inputs from Land into the Ocean. *Science*, 347(6223), 768–771. <https://doi.org/10.1126/science.1260352>
- [12] Li, T. M. (2014). *Land's End: Capitalist Relations on an Indigenous Frontier*. Durham: Duke University Press.
- [13] Merchant, C. (1980). *The Death of Nature: Women, Ecology, and the Scientific Revolution*. New York: Harper & Row.
- [14] Naess, A. (1973). The Shallow and the Deep, Long-Range Ecology Movement. A Summary. *Inquiry*, 16(1–4), 95–100. <https://doi.org/10.1080/00201747308601682>
- [15] Plumwood, V. (2002). *Environmental Culture: The Ecological Crisis of Reason*. London: Routledge.
- [16] Rangkuti, N. (2009). *Tambang, Konflik, dan Masyarakat Adat* [Mining, Conflict, and Indigenous Communities]. Jakarta: ELSAM.
- [17] Rasiah, R., Wantasen, I. L., Tulung, G. J., Rauf, R., Bilu, L., & Sofian, N. I. (2024). 'I Have Been Faithful to Thee, Cynara': Parodying Canonical Pro-Slavery Novel in The Wind Done Gone. *World Journal of English Language*, 14(3), 62. <https://doi.org/10.5430/wjel.v14n3p62>
- [18] Rauf, R., Junaid, S., Ulya, A., Panjaitan, I. P., Widhyah Saputri, D. Y., & Lintang Gritantin, L. A. (2024a). How Does the Buginese Woman Image in Ketika Saatnya? (Postcolonial Studies and Islamic Perspective). *Elite: English and Literature Journal*, 11(2), 231–244.
- [19] Rauf, R., Magfirah, S., Ridwan, R., & Ulya, A. (2024b). What Happened in Talaga Biru? A Study of Ecological Literature. *Humanitatis: Journal of Language and Literature*, 10(2), 301–312. <https://doi.org/10.30812/humanitatis.v10i2.4079>
- [20] Rauf, R., Muslimin, M. F., Ulya, A., & Gritantin, L. A. L. (2025). Discourse on gender: A Linguistic Analysis of Body Autonomy and Patriarchal Narratives in Wa Ndiu-Diu. *Studies in English Language and Education*, 12(1), 521–534. <https://doi.org/10.24815/siele.v12i1.38267>
- [21] Shiva, V. (2005a). *Earth Democracy: Justice, Sustainability, and Peace*. New York: South End Press.
- [22] Shiva, V. (2005b). *Ecofeminisme: Perspektif Gerakan Perempuan & Lingkungan* (Diterjemahkan oleh Kelik Ismunanto dan Lilik) [Ecofeminism: Women's Movement & Environmental (translated by Kelik Ismunanto and Lilik)]. Yogyakarta: IRE Press.
- [23] Subuh, R. D., Rauf, R., Maricar, F., & Ulya, A. (2024). The Utilization of Ternate Oral Tradition for the Identification of Potential Disasters. *Etnorefika: Jurnal Sosial dan Budaya*, 13(1), 134–143.
- [24] Taha, M., Febriningsih, F., Asfar, D. A., Zalmansyah, A., Muzammil, A. R., Ajam, A., Lestari, S. A. B., Zar'in, F., Rauf, R., & Gritantin, L. A. L. (2024). Sula Malay's Hate Speech in the Sanana Jurisdiction: A Pragmatic Study. *Journal of Language Teaching and Research*, 15(6), 1941–1950. <https://doi.org/10.17507/jltr.1506.19>
- [25] United Nations Environment Programme. (2021). *From Pollution to Solution: A Global Assessment of Marine Litter and Plastic Pollution*. Kenya: UNEP.
- [26] United Nations Environment Programme. (2022). *Global Environment Outlook 6: Healthy Planet, Healthy People*. Kenya: UNEP.
- [27] World Bank. (2021). *The Economics of Biodiversity: The Dasgupta Review*. London: HM Treasury.



Wildan has a Ph.D. in literature. He is a lecturer in the Indonesian Literature Study Program at Universitas Khairun. His teaching and research focus on literary research methods, ecological literature, and oral literature. wildan@unkhair.ac.id



Ramis Rauf has a Ph.D. in literature. He is a lecturer in the Indonesian Literature Study Program at Universitas Khairun. His teaching and research focus on Indonesian literature, linguistics, and cultural studies. He has been actively involved in exploring the role of local literature in shaping cultural identity and strengthening national character. His recent works examine the intersection between language, literature, and society. ramis.rauf@unkhair.ac.id



Ratna is a lecturer in the Indonesian Literature Study Program at Universitas Khairun. Her teaching and research focus on literacy. ratna@unkhair.ac.id



Afriani Ulya is a lecturer at the Indonesian Literature Study Program, Universitas Halu Oleo. Her teaching and research focus on psychoanalysis, sociology of literature, and gender. Her scholarly interests include Indonesian literature, language education, and the preservation of local cultural heritage. She has contributed to studies on integrating literature in educational curricula and promoting regional literary traditions. Her research also addresses the role of literature in fostering intercultural understanding. afrianiulya11494@uho.ac.id



Vidhia Agmareina Hirto is a lecturer in the Travel and Tourism Business Program at Universitas Khairun. She currently serves as the Head of the Travel Business Teaching Factory. She holds a Master's degree in Tourism Studies from Universitas Gadjah Mada. vidhiahirto@unkhair.ac.id



Sulmi Magfirah is a lecturer at the English Literature Study Program, Universitas Khairun. Her teaching and research focus on English literature, children's literature, and comparative literature. She has written on identity, narrative, and intercultural perspectives in literature. Her current research interests include the role of literature in fostering critical thinking and cross-cultural awareness among students. sulmi@unkhair.ac.id



Fitriningsih Pratiwi Mahmud was born in Kaimana on 15 November 1992. She serves as a Lecturer in the English Literature Department at Universitas Khairun. Beyond her academic role, her active participation in foundations such as The Tebings and Sabua Foundation highlights her commitment to preserving North Maluku's cultural heritage and empowering local communities. fitriningsih.pratiwi@gmail.com



Dhini Yustia Widhyah Saputri was born in Ngawi on 5 September 1990. She completed elementary education in East Halmahera, North Maluku. She continued her undergraduate studies in the Study Program of Indonesian Literature at Sebelas Maret University, Surakarta, and her Master's in Literature at Gadjah Mada University. She has had an interest in philology and classical literature since her undergraduate studies and continues to pursue them to this day. Currently, she is a permanent lecturer in the Indonesian Literature study program at Khairun University, Ternate, North Maluku. dhini.yustia@unkhair.ac.id



Ulfah Dwi Januarty was born in Bitung on 27 January, 1989. She is a lecturer in the Department of Indonesian Literature at Universitas Khairun. She earned her Master's degree in Literary Studies from Universitas Diponegoro, where she focused on literature and contemporary literary discourse. Her teaching portfolio includes courses in creative reading, essay writing, and Indonesian literary history. ulfah@unkhair.ac.id



Ahmad Rifani Talaohu was born in Ambon on 29 December 1991. He is a Lecturer in the English Department at Universitas Khairun. He completed a Master's degree in General and Applied Linguistics at the Australian National University and is currently pursuing a Ph.D. in Linguistics at the University of Newcastle, Australia. He is an awardee of LPDP Scholarship. ahmadrifani@unkhair.ac.id

Exploring Lexical Processing and Sustainability Awareness in EFL Learners: A Psycholinguistic Discourse-Cognition Perspective

Mohammad Abdulaal*

Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Kingdom of Saudi Arabia

Ahmad Sharaf Eldin

Department of English Language, Faculty of Arts, Menoufia University, Egypt;
Educational Consultant, Ministry of Defence, Abu Dhabi, UAE

Hanan Al-Johani

Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Kingdom of Saudi Arabia

Mohammad Aldawooddy

Department of English, Faculty of Arts, Port Said University, Port Said, Egypt

Abstract—Despite the increasing emphasis on sustainability literature, limited research has examined how EFL learners cognitively and neurologically process sustainability-related literature in English. This research study investigates this psycholinguistic gap by analyzing how EFL learners are engaged in word processing while comprehending sustainability discourse and developing awareness of worldwide environmental issues. The research endeavors to answer the following question: How do psycholinguistic processes, especially word processing, impact EFL learners' comprehension and sustainability insights in discourse-intensive contexts? A mixed-method design was used, integrating quantitative visual tracking indicators of word recognition and processing with qualitative reflective narrations and open-ended interviews. Forty intermediate EFL university learners participated in this research study. Data analysis focused on processing cognitive pace, precision, and patterns of cognitive engagement, alongside learners' reported comprehension of sustainability notions. Findings show that learners' word processing effectiveness significantly correlates with their full comprehension of sustainability-based discourse. Participants who showed a higher level of lexical and semantic processing demonstrated greater awareness of ecological issues. Further, they reflected a more precise interpretation of discourse meaning. These results suggest that some psycholinguistic factors, such as word recognition and semantic integration, play an essential part in linking language acquisition to content-based insights. The research implications include integrating targeted psycholinguistic activities that promote word processing skills while simultaneously boosting learners' sustainability insights.

Index Terms—word processing, cognitive linguistics, language perception, sustainability texts, discourse

I. INTRODUCTION

Incorporating sustainability discourse into scholastic frameworks has become essential as worldwide ecological challenges augment. Within the EFL context, this integration boosts language proficiency. Furthermore, it fosters environmental perspectives among language users. However, the cognitive processes underlying learners' engagement with sustainability texts are still unexplored. These cognitive processes receive scant research attention. Understanding how English language learners process this discourse is crucial for designing effective scholastic mechanisms that boost both linguistic and environmental literacy (Lin & Lei, 2020).

Psycholinguistics extensively examined cognitive word processing in first and second language acquisition, concentrating on lexical access, semantic integration, and reading perception. Psycholinguistic studies have provided insights into how language users process language at the cognitive levels. However, there are relatively few studies that explore how EFL learners and language users comprehend sustainability discourse. While some studies have investigated the incorporation of environmental themes into second language courses (Chen, 2021), the cognitive mechanisms and operations engaged in processing such content have not been adequately examined. This linguistic gap is substantial, as understanding the cognitive mechanisms involved can reveal more effective instructional approaches that boost both

* Corresponding Author.

language skills and environmental awareness (Rayner et al., 2006).

This study attempts to examine this gap by investigating how language users process sustainability-themed texts and how this processing reflects their environmental awareness. By using eye-tracking technology to estimate real-time word processing and combining it with qualitative think-aloud protocols, this research endeavors to provide a comprehensive analysis of the cognitive mechanisms at play. The findings are anticipated to offer valuable insights into the interplay between language processing and content perception in second language acquisition contexts (Odia, 2021; Nervino et al., 2024).

Guided by the identified gaps and objectives, this study seeks to explore several key questions related to English language learners' engagement with sustainability discourse. Specifically, it examines how learners process sustainability-themed texts at both the lexical and semantic levels, investigates the relationship between their word processing efficiency and their overall comprehension of sustainability discourse, and analyzes how cognitive processing of such texts influences learners' environmental awareness.

II. LITERATURE REVIEW

The intersection of psycholinguistics, English language learning, and sustainability awareness is an emerging area of research that seeks to understand how cognitive processes impact learners' engagement with ecological discourse. This literature review synthesizes existing research studies on word processing in second language acquisition contexts, the role of psycholinguistics in language learning, and the integration of sustainability themes into language acquisition (Al-Khotaba et al., 2019).

Research on second language (L2) reading has consistently concentrated on the centrality of psycholinguistic cognitive operations in facilitating text perception. Word recognition and lexical access are considered crucial, with efficient processing enabling language users to handle textual semantics more effectively (Perfetti & Stafura, 2014). Eye-tracking research studies further demonstrated that measures such as fixation durations and regressions offer critical insights into the cognitive operations underlying cognitive processing (Rayner, 2009). These findings show that reading is not a passive decoding process but rather a highly active, cognitively demanding activity that incorporates multiple linguistic and semantic markers.

Many reading models, such as Beers (1987), Geyer (1972), and Schraw and Bruning (1996), propose that up to five words can be processed simultaneously. While prior research indicates that foveal words may be processed alongside immediately adjacent words, it remains unclear whether this parallel processing extends to three words. To investigate this point, Mlinarič et al. (2025) employed a lexical decision paradigm that presents three flankers on each side of a central target. Results showed that only flankers directly adjacent to the target affected word recognition, whereas more distant flankers had no measurable or significant influence. These findings indicate that readers can process approximately three words in parallel. This level of parallel processing exceeds the predictions of some models, such as the E-Z Reader and the Chinese Reading Model, but it remains lower than the estimates proposed by other frameworks, including Glenmore and OB1-Reader (Mlinarič et al., 2025; Snell & Grainger, 2018).

In EFL contexts, difficulties at the lexical and semantic levels often impede learners' comprehension of academic or content-based texts (Koda, 2005; Nassaji, 2011). Vocabulary knowledge, including both breadth and depth, is a strong predictor of reading performance, particularly when learners engage with dense and conceptually demanding discourse (Laufer & Ravenhorst-Kalovski, 2010). Content and language integrated learning (CLIL) approaches emphasize the importance of embedding linguistic instruction within meaningful subject matter, thereby supporting both language and conceptual development (Coyle et al., 2010). Yet, despite the rising global importance of sustainability, studies that directly examine learners' cognitive engagement with sustainability texts remain limited (Sadoughvanini & Ghaemi, 2025).

Gahshan and Weintraub (2024) argued that elementary-school students are increasingly expected to compose texts using computers, which requires both higher-level skills, such as planning, translating, and revising, and lower-level transcription skills. To address this issue, the Word Processing and Typing Instructional Program (WoTIP) was developed and implemented through a collaborative consultation model. This study examined whether WoTIP could improve typing speed, word processing, and composition performance compared to a control group with no typing or word processing instruction. The findings indicate that WoTIP is an effective approach for enhancing both transcription skills and overall writing abilities in elementary students.

Research studies in the sustainability context underscore its transformative potential in boosting critical thinking and global awareness. Sterling (2010) points out that sustainability pedagogy fosters learners' capacity for reflection and systemic thinking, while Chun (2016) underscores the role of language processing in promoting sustainability consciousness. Integrating sustainability discourse into English language curricula has been shown to contribute not only to linguistic development but also to learners' environmental and ecological awareness and critical engagement with global concerns. However, few studies connect these outcomes to psycholinguistic processes such as word recognition and semantic integration, leaving a gap in the literature that this study seeks to address (Piga, 2023).

Prior psycholinguistic research in Greek and English has demonstrated that the lexical category and argument structure of the base verb play a crucial role in processing deverbal pseudowords that violate these constraints. A comparable study in Slovenian highlighted the influence of aspectual properties of the base verb on deverbal word processing. Tsaprouni

and Manouilidou (2025) revisit the impact of aspect on the morphological processing of Greek deverbal formations. To examine this, both an offline acceptability judgment task and an online lexical decision task were administered using various pseudowords that breached derivational rules. Findings indicate that aspect significantly influences the processing of deverbal pseudowords in Greek. Specifically, aspectual rules interact distinctly with categorial and argument structure constraints for the suffix responsible for generating formations with unambiguous, eventive interpretations and clear eventive implications, but this effect does not extend to other suffixes.

This study is grounded in psycholinguistic theoretical frameworks of reading comprehension. The Lexical Quality Hypothesis (Perfetti, 2007) posits that efficient comprehension relies on high-quality lexical representations that incorporate form, meaning, and context. In the context of sustainability-based discourse, where language users often encounter specialized vocabulary, lexical quality plays a substantial role in determining processing efficiency and overall language comprehension (Wang, 2023; Bilić Meštrić & Šimičić, 2017).

The Construction-Integration Model, put forth by Kintsch (1998), defines language perception as a dynamic operation in which readers develop the semantic content from linguistic input and integrate it with prior knowledge structures. This framework is pertinent to sustainability texts, which require language users to connect new linguistic lexical items with broader prior knowledge about ecological concerns (Yang & Peng, 2022).

Furthermore, this research relies on a content-based language learning theoretical framework. It stipulates that language learning is most significant when embedded in meaningful, content-based contexts. Sustainability texts provide authentic domains in which language users can simultaneously advance their language proficiency and develop awareness of global challenges. These theoretical perspectives provide a sound basis for investigating how psycholinguistic operations—specifically word recognition, semantic integration, and cognitive engagement—contribute to both comprehension and sustainability awareness in EFL settings (Coyle et al., 2010; Eriksson, 2015).

III. RESEARCH METHODS

This research study used a mixed-methods research design to examine how English language learners cognitively process sustainability-related discourse and how psycholinguistic mechanisms of word processing shape comprehension and awareness of ecological issues. The integration of quantitative and qualitative techniques was deemed essential, as it paved the way for a comprehensive exploration of both measurable processing patterns and learners' subjective experiences. While quantitative data provided insights into lexical and semantic processing efficiency, qualitative data reflected the depth of learners' involvement with sustainability terminologies and their reflective awareness of environmental discourse.

A. Research Design

A convergent mixed-methods design was used, incorporating eye-tracking experiments with think-aloud protocols in addition to semi-structured interviews. This research design ensured that quantitative measures of word recognition, reading speed, and precision were triangulated with qualitative descriptive accounts of subjects' interpretive operations. The rationale for choosing this design was to spotlight the cognitive factors of reading sustainability discourse while also considering subjects' meaning-developing strategies. The mixed-methods design facilitated a broader understanding of the correlation between psycholinguistic processing and the perception of content-rich texts, bridging cognitive data with learner-reported experiences.

B. Participants

Forty intermediate-level university students participated in this study. The participants were enrolled in English language courses at the time of data collection. Participants were chosen through purposive sampling to ensure relative homogeneity in language proficiency, thus reducing the impact of wide-ranging linguistic competence on the study's findings. Proficiency was ensured through institutional placement tests. All subjects gave their informed consent. In addition, ethical approval was obtained from the host institution's research committee. The balance of gender distribution was observed. The subjects represented diverse academic majors, ensuring exposure to sustainability topics across different disciplinary environments. All participants have been chosen according to standardized placement test scores.

C. Instruments

Three primary instruments were employed. First, an eye-tracking tool (Tobii Pro X3-120 eye-tracker; 120 Hz sampling frequency) was used to obtain quantitative measures of lexical and semantic processing, focusing on word recognition speed, fixation durations, and regressions while reading sustainability-related texts. Eye-tracking enabled real-time observation of participants' attentional shifts and processing efficiency. Second, think-aloud protocols were employed to document participants' cognitive strategies during reading. Participants verbalized their thoughts, interpretations, and difficulties as they engaged with the texts, providing qualitative data on metacognitive engagement and semantic integration. Third, semi-structured interviews were conducted immediately after the reading sessions to further explore learners' understanding of sustainability terminology and their attitudes toward environmental discourse. Sustainability awareness was assessed through these latter two instruments, together with comprehension-oriented questionnaire items, allowing awareness to be evaluated independently of eye-tracking measures. The triangulation of these methods ensured

both depth and breadth in data collection and enabled a comprehensive assessment of processing patterns and reflective awareness.

D. Materials

The reading resources consisted of carefully chosen sustainability-themed texts adapted from authentic and journalistic sources. These texts covered global environmental topics such as climate change and sustainable consumption. Texts were piloted with a small group of participants. The aim was to make sure that they were appropriate in terms of lexical length and textual transparency. The final set of texts balanced between academic vocabulary and comprehensible discourse, thereby aligning with the intermediate proficiency level of the participants. Furthermore, key sustainability terminologies were incorporated in the texts to facilitate the process of observing lexical and semantic processing.

E. Procedures

The process of collecting data was conducted in three stages. In the first stage, the study participants conducted an eye-tracking experimental test. Each participant sat in front of a computer equipped with an eye-tracking device. This device records both fixation patterns and gaze duration while reading the sustainability-themed texts. Each session spanned 20 minutes. In the second phase, each subject was instructed to complete a think-aloud protocol while reading some texts. Prior to this phase, the subjects were trained in the think-aloud technique to reduce interruptions in cognitive processing. Their verbalizations were audio-recorded for later transcription and analysis. In the third stage, semi-structured interviews were conducted within one week of the reading tasks. Interviews investigated learners' comprehension of sustainability discourse, perception of ecological issues, and attitudes towards how word-level processing impacted their understanding. Each interview lasted between 25 and 40 minutes.

F. Data Analysis

Quantitative numerical data from the eye-tracking sessions were investigated using inferential and statistical data. Metrics such as fixation time, regression frequency, and overall reading durations were computed to determine lexical and semantic processing efficiency. Correlational analyses were conducted to explore the relationship between reading processing rate and comprehension scores. In the qualitative analysis, think-aloud protocols, sometimes called concurrent verbalizations, and interview data were investigated through thematic coding. Coding schemes concentrated on lexical decoding mechanisms, semantic integration, interpretive challenges, and expressed awareness of sustainability topics. To verify reliability, two investigators independently coded the dataset. They addressed inconsistencies through discussion. Triangulation across data sources boosted the validity of the results by comparing data from eye-tracking, verbal protocols, and interviews.

IV. RESEARCH RESULTS

This section presents the results of the research study by integrating quantitative and qualitative analyses. Findings are framed around the three key foci. The first focal point is lexical and semantic processing of sustainability texts. The second one is the correlation between word processing effectiveness and perception. The third is the relationship between cognitive processes and environmental awareness. Quantitative and inferential statistics are presented alongside qualitative insights from think-aloud tests and interviews. Statistical assessments were selected based on the data collected and the research questions. Eye-tracking tests revealed that the participants had allocated more cognitive effort to eco-themed vocabulary than to general lexical items. The average fixation duration for sustainability terms was longer, and regression counts were higher, suggesting greater processing difficulty. Comprehension scores varied, with higher-proficiency learners showing stronger outcomes.

TABLE 1
DESCRIPTIVE STATISTICS FOR LEXICAL PROCESSING AND PERCEPTION

Measure	Mean (M)	Standard Deviation (SD)	Range	Notes
Fixation duration (sustainability words, ms)	274	35.7	220–340	Longer fixations for domain-specific words
Fixation duration (general words, ms)	248	30.4	200–310	Shorter than sustainability words
Regression counts (sustainability words)	4.12	1.8	2–8	More regressions were observed.
Regression counts (general words)	2.75	1.2	1–5	Fewer regressions were observed.
Comprehension scores (%)	72.6	8.9	58–89	Higher proficiency was linked to better comprehension.

A paired-samples t-test showed significantly longer fixation times for eco-related words than for general lexical terms. It is due to the fact that $t(39) = 6.47$ and $p < .001$. A Wilcoxon signed-rank test confirmed significantly higher regression metrics for sustainability words. It is owing to the fact that $Z = -4.28$ and $p < .001$. These findings showed that specialized lexical items required more cognitive processing.

Correlational analyses showed that processing efficiency was significantly associated with learners' perception of sustainability-related discourse. Specifically, fixation duration was found to correlate negatively with comprehension as $r = -.46$ and $p = .004$. It indicated that participants who spent less time fixating on key terms achieved higher comprehension scores. By the same token, regression metrics were also negatively correlated with perception, $r = -.42$ and $p = .006$. It indicated that frequent rereading was linked to lower comprehension outcomes. In contrast, total reading time revealed no significant relationship with comprehension, $r = -.12$ and $p = .41$. It revealed that efficiency at the word-processing level, rather than overall time spent, was the critical factor affecting comprehension performance.

A multiple regression model confirmed that both fixation duration and regression counts significantly predicted comprehension scores, $F(2, 37) = 7.81$, $p = .001$, explaining 29% of the variance.

TABLE 2
CORRELATION AND REGRESSION RESULTS FOR PROCESSING EFFICIENCY AND COMPREHENSION

Predictor Variable	Correlation with Comprehension	Regression Coefficient (β)	Significance (p)
Fixation duration (ms)	$r = -.46^{**}$	-0.34	.01
Regression counts	$\rho = -.42^{**}$	-0.31	.02
Total reading time (ms)	$r = -.12$ (n.s.)	—	—

Note: $p < .05$; ** indicates significant correlation.

An independent sample t-test revealed that higher-intermediate participants significantly outperformed lower-intermediate ones in comprehension scores. It is due to the fact that $t(38) = 4.12$ and $p < .001$. Eye-tracking data also showed shorter fixation durations and fewer regressions for the higher-intermediate group. It reflected greater lexical efficiency.

TABLE 3
GROUP COMPARISONS BY PROFICIENCY LEVEL

Measure	Higher-Intermediate	Lower-Intermediate	t-value	p-value
Comprehension scores (%)	80.2 (SD = 6.7)	69.4 (SD = 7.1)	4.12	< .001
Fixation duration (sustainability, ms)	259 (SD = 24.1)	288 (SD = 31.6)	-3.51	.001
Regression counts (sustainability words)	3.45 (SD = 1.5)	4.79 (SD = 1.9)	-2.97	.005

Awareness levels, coded from interviews, were significantly associated with processing efficiency. A chi-square test of independence showed that learners with high processing efficiency were more likely to demonstrate high awareness of sustainability issues. It is due to the fact that $\chi^2(2, N = 40) = 8.73$ and $p = .013$.

The analysis of verbal reports showed four key themes that threw light on the differences between high- and low-efficiency learners. First, lexical inference strategies played a substantial role in comprehension. Participants with higher processing efficiency were able to make faster and more precise inferences about the meanings of sustainability-related lexical items. On the other hand, lower-efficiency learners often hesitated or relied on repeated attempts to decode meaning. Second, semantic incorporation is considered an important factor, with successful comprehension closely tied to the learners' ability to incorporate domain-specific lexical items into the broader discourse of the texts. High-efficiency participants showed greater skill in linking core terms such as "renewable energy" or "carbon footprint" to the key arguments of the text. Low-efficiency learners, on the other hand, often treated these terms in isolation, limiting their understanding.

The third and fourth themes highlighted the metacognitive and reflective dimensions of comprehension. In terms of metacognitive monitoring, high-efficiency learners verbalized fewer challenges while reading. It indicated more automatic and fluent processing. Low-efficiency participants frequently expressed confusion about articulate comprehension struggles. Finally, the topic of awareness structuring highlighted a significant distinction between the two groups. Participants with efficient processing were more likely to make meaningful connections between the thematic content of the texts and wider global environmental issues. It reflected an ability to extend language learning into real-world awareness. Conversely, low-efficiency learners tended to remain focused on decoding vocabulary items without linking them to broader sustainability contexts. It resulted in a narrower and less integrated understanding.

Interview data showed that high-efficiency participants provided precise and cognitively detailed definitions of sustainability terminologies. On the other hand, low-efficiency learners gave vague or incomplete explanations. High-efficiency learners also pointed out that engaging with the texts broadened their perspective on environmental issues. Furthermore, it linked their language skills with global awareness.

The convergence of findings supports the argument that psycholinguistic efficiency underpins both comprehension and awareness. Efficient processors displayed measurable lexical advantages, higher comprehension scores, and greater sustainability consciousness. Conversely, low-efficiency participants illustrated cognitive load and restricted awareness.

The findings of the research study showed that sustainability-related lexical items demanded a greater processing load in comparison to general vocabulary. It reflected the cognitive load involved in engaging with domain-specific terms. Importantly, efficient word processing was found to be a strong predictor of perception performance. It is due to the fact that learners who showed shorter fixation durations and fewer regressions achieved higher comprehension scores. This

pattern was particularly clear among higher proficiency participants, who processed sustainability texts more effectively and showed superior overall perception outcomes.

Beyond comprehension, processing efficiency was also closely tied to learners' broader understanding and awareness of environmental issues. Learners with more efficient processing not only displayed stronger sustainability awareness but also showed greater ability to connect linguistic input with meaningful content knowledge. The qualitative findings revealed that efficient processing facilitated deeper semantic integration. Further, it encouraged learners to reflect critically on global environmental issues. These results suggest that psycholinguistic efficiency serves as a bridge between language learning and the growth of sustainability consciousness.

V. DISCUSSION

This research study investigated how English language learners process eco-discourse at both the lexical and semantic levels. Furthermore, it examined how cognitive linguistic efficiency affects both language perception and environmental consciousness. This study used a qualitative-quantitative design. It combined visual tracking (i.e., focus on fixation patterns), verbalizations, and semi-structured interviews. The results underscored the decisive role of word processing in linking language learning with the development of sustainability awareness. The discussion below highlights the points of agreement, divergence, and contribution to existing literature.

One of the key findings of the study was that sustainability lexical items required significantly greater cognitive effort than general vocabulary. It was indicated by extended fixation durations and higher regression metrics. This is in line with prior psycholinguistic research showing that low-frequency lexical items generally place a greater cognitive load on readers in comparison to high-frequency words (Kuperman et al., 2012; Rayner, 2009; Jin & Li, 2025). Empirical research in L2 reading has also highlighted that learners give more attention to unfamiliar or semantically dense lexical items, which in turn slows down overall processing (Chaffin et al., 2001). The present findings extend this body of work by affirming that sustainability discourse, much like other specialized domains such as medical or legal English, poses unique challenges for non-native learners of English (Anon, 2016; Hilterbran, 2016).

At the same time, the results diverge from studies suggesting that repeated exposure to thematic vocabulary can quickly reduce processing load (Schmitt, 2010; Pellicer-Sánchez, 2016). In the present research study, despite participants being intermediate-level learners who may have had prior exposure to some sustainability terms, the difficulty persisted. One possible explanation is the dual challenge posed by sustainability discourse: not only does it include low-frequency terminologies, but it also includes abstract and globally oriented terms that require higher-order integration. Hence, the findings revealed that sustainability texts presented a double layer of difficulty and conceptual complexity.

Another significant outcome of this research study was the strong predictive correlation between efficient word processing and comprehension performance. Shorter fixation intervals and lower regression are inversely correlated with comprehension complexities. Regression analysis affirmed that processing efficiency explained nearly 33% of the variance in comprehension outcomes. These results are consistent with the eye-mind model (Just & Carpenter, 1980). This model posits that the time spent fixating on a lexical item reflects the cognitive processing required to perceive it. Therefore, high-proficiency readers spend less time on word-level deciphering and can assign more processing power to higher-level semantic construction.

This finding is consistent with the work of Grabe (2009), who pointed out that fluent word recognition is crucial to perception in L2 reading. Similar conclusions have been reported in empirical studies by Koda (2005) and Yamashita (2002), both of whom found that lexical processing speed is one of the strongest predictors of comprehension among EFL and ESL learners. In contrast, some research suggests that extended reading time can sometimes enhance comprehension because it allows readers to engage in deeper elaboration (Anderson, 2009). However, the current study's results suggest that in the case of sustainability discourse, efficiency at the word level, rather than total time spent, is the critical factor. This discrepancy may be explained by the content-specific nature of the texts: since sustainability discourse is conceptually demanding, slow word-level processing may overload working memory, leaving fewer resources available for semantic integration.

The study also evidenced distinct variations between higher- and lower-intermediate learners. Higher proficiency participants demonstrated shorter fixation durations and fewer regressions. They also showed superior comprehension scores compared to the lower proficiency participants. These findings are well-aligned with L2 acquisition research. Indeed, second language research repeatedly emphasizes that proficiency level moderates cognitive processing efficiency (Alderson, 2000; Grabe, 2009). Higher proficiency learners usually have larger lexical items, more robust decoding skills, and more developed background knowledge. All these skills facilitate efficient reading.

This finding is in line with the results of Yamashita and Shiotsu (2017), who found that higher proficiency learners exhibit faster lexical access and greater automaticity in word recognition, both of which directly support comprehension. Similarly, Jeon and Yamashita's (2014) meta-analysis of L2 reading research concluded that language proficiency explains a significant proportion of variance in comprehension outcomes across diverse populations. The current study reinforces this conclusion in the specific context of sustainability discourse, emphasizing that lexical efficiency is not an isolated skill but one intertwined with overall language competence.

Nevertheless, the results contrast with studies suggesting that even lower proficiency learners can achieve comprehension with adequate scaffolding and use of strategies (Anderson, 2009; Zhang & Wu, 2009). In the present

study, think-aloud protocols (e.g., verbalizing word meanings while reading sustainability texts) revealed that lower proficiency learners often attempted to use contextual inference, but these strategies were less effective due to their limited lexical and syntactic knowledge. This suggests that while strategies may compensate for gaps in knowledge in less complex texts, the high cognitive load of sustainability discourse magnifies the differences between proficiency groups.

A highly significant contribution of this study resides in demonstrating the correlation between processing efficiency and sustainability awareness. Participants with high efficiency were significantly more likely to display strong awareness of sustainability concepts in interviews. Lower-efficiency learners often encountered difficulties articulating precise or meaningful interpretations. This suggests that psycholinguistic processing promotes the transformation of textual understanding into metacognitive insights.

The finding corresponds with prior research in content-based language learning, which emphasizes that language skills and disciplinary knowledge are mutually reinforcing (Coyle et al., 2010). In studies of CLIL (Content and Language Integrated Learning), Dalton-Puffer (2011) similarly observed that learners with stronger language processing abilities were better positioned to develop subject-matter understanding. In the context of sustainability, the present findings align with Sterling (2010), who contends that sustainability education demands more than mere knowledge acquisition; it requires equipping EFL learners with the cognitive and linguistic skills needed to engage with complex concepts.

However, this study diverges slightly from some studies on awareness development, such as those by Alptekin and Erçetin (2010), which suggested that even when comprehension is incomplete, learners may still develop partial conceptual awareness. In the present study, learners with inefficient processing did not demonstrate substantial awareness, possibly because sustainability discourse requires precise integration of terminology and concepts to form accurate understandings. Unlike narratives or general expository texts, sustainability texts rely heavily on specific lexical items as carriers of meaning, and failure to process these efficiently appears to hinder awareness development.

The qualitative outcomes from think-aloud protocols and interviews enriched the quantitative results by highlighting how efficient and inefficient learners approached sustainability discourse. Cognitively efficient learners exhibited effective lexical inference strategies. Further, they showed effortless semantic integration in addition to greater metacognitive reflection. Conversely, low-efficiency learners exhibited hesitation and frequent rereading. They also displayed a narrow concentration on isolated words. These findings are in alignment with studies on reading strategies in EFL contexts. These studies stress that proficient learners are engaged in more meaning-driven processing. That is, they connect words to overall textual discourse (Block, 1992; Anderson, 2009).

Moreover, the results support the metacognitive reading model proposed by Baker and Brown (1984). This model argues that successful comprehension is associated with the ability to monitor understanding and regulate cognitive resources. High-efficiency participants in this study verbalized fewer complexities. It suggested that they had internalized more automatic decoding skills. Such abilities enabled them to focus on monitoring meaning construction. This corresponds to the findings from Zhang (2001). In fact, Zhang (2001) pointed out that efficient L2 readers exhibit higher levels of metacognitive consciousness. Weaker readers, on the other hand, concentrate on lexical decoding.

The qualitative outcomes also enhance the broader literature on sustainability discourse. Sterling (2010) and Wals (2011) argue that sustainability awareness requires learners to connect text-based input to real-world issues through reflection and critical thinking. The current study shows that such connections are more likely to occur when learners are not cognitively overloaded by basic decoding, further underlining the pedagogical importance of psycholinguistic efficiency.

Taken together, the findings of this study are largely consistent with existing research on second language reading and content-based learning, but they also extend the literature by explicitly linking psycholinguistic processing to sustainability awareness. Evidence that efficient word processing predicts comprehension and awareness complements longstanding findings in reading research while applying them to a novel domain.

This integrative perspective corresponds to the view of language learning as both a cognitive and a sociocultural process (Lantolf, 2006). On the cognitive side, the results confirm that efficient processing mechanisms are critical for comprehension. On the sociocultural side, the outcomes show that when learners can process language fluently, they are better able to engage with the content in a way that contributes to their global and environmental awareness. Few previous studies have directly examined this dual role of psycholinguistic efficiency. It makes this study a significant addition to the interface of language learning and sustainability consciousness.

While most of the findings reinforce previous literature, some divergences highlight new insights. For example, the finding that total reading time did not predict comprehension diverges from research suggesting that longer engagement can support deeper understanding (Anderson, 2009). The present results suggest that in cognitively demanding domains like sustainability, efficiency rather than sheer effort is what drives comprehension. Similarly, the finding that low-efficiency learners did not gain substantial sustainability awareness contrasts with studies suggesting that partial comprehension can still yield partial content learning (Alptekin & Erçetin, 2010; Chamusca, 2005). This indicates that sustainability discourse may demand a threshold level of lexical and semantic precision for awareness to emerge.

VI. CONCLUSION AND IMPLICATIONS

This research study immensely contributes to the growing literature on psycholinguistics and content-based instruction. It demonstrates that word processing efficiency plays a central role in EFL learners' perception and awareness of

sustainability discourse. Findings confirm that sustainability-related vocabulary demands more cognitive effort. Efficient processing predicts comprehension outcomes. Highly proficient learners are more successful in integrating text meaning. Efficiency is closely associated with sustainability awareness. These outcomes are consistent with prior research on reading and content learning. These findings also extend the field by explicitly linking psycholinguistic mechanisms with global environmental consciousness.

The results reveal that efficient psycholinguistic processing functions as a link between language proficiency and the growth of ecological consciousness. This outcome sheds light on the value of incorporating explicit vocabulary instruction for sustainability terminology and using diagnostic tools to assess lexical processing difficulties. It also implies that improving lexical and semantic efficiency should be a key focus of instruction. Educators can equip learners with tools to process sustainability discourse effectively. It can contribute to language development and to fostering learners' ability to engage critically with one of the most pressing global issues of the 21st century.

The results carry important implications for EFL learning and acquisition. First, the finding that sustainability vocabulary requires greater cognitive effort implies the need for explicit instruction in domain-specific lexis. This can be performed through pre-teaching key terms or using vocabulary notebooks. Second, as long as processing efficiency predicts language perception, learning interventions should concentrate on enhancing automaticity in word processing. It can be attained through fluency-building drills and constant reading. Third, the link between efficiency and awareness suggests that integrating sustainability discourse into language curricula can develop linguistic proficiency. Further, it can boost world citizenship.

These instructional implications correspond to content-based pedagogical frameworks. These instructional approaches show that teaching language via meaningful content fosters language outcomes. When incorporating sustainability texts into English language courses, instructors can achieve the linguistic objectives. In addition, they can engage students with ecological concerns.

While the research study gives robust findings, it also has some limitations. The sample size of forty participants limits generalizability. Future studies could include more diverse populations. Thus, these studies can test whether the results hold across different proficiency levels and cultural contexts. Additionally, the study concentrated only on intermediate-level learners. Research on advanced or beginner learners could reveal different processing dynamics.

ACKNOWLEDGEMENTS

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/02/34187).

REFERENCES

- [1] Al-Khotaba, H., Alkhataba, H., Abdul-Hamid, S., & Bashir, I. (2019). Foreign Language Speaking Anxiety: A Psycholinguistic Barrier Affecting Speaking Achievement of Saudi EFL Learners. *Arab World English Journal*, 10(4), 313–329. <https://doi.org/10.24093/awej/vol10no4.23>
- [2] Anon. (2016). The existence and the sustainability of the Tantrism in Bali. *International Journal of Linguistics, Literature and Culture*. <https://doi.org/10.21744/ijllc.v2i4.286>
- [3] Beers, T. (1987). Commentary: Schema-Theoretic Models of Reading: Humanizing the Machine. *Reading Research Quarterly*, 22(3), 369. <https://doi.org/10.2307/747974>
- [4] Billić Meštrić, K., & Šimičić, L. (2017). Language orientations and the sustainability of Arbanasi language in Croatia – A case of linguistic injustice. *Open Linguistics*, 3(1). <https://doi.org/10.1515/opli-2017-0008>
- [5] Chamusca, I. L., Pagano, T. P., & Loureiro, R. B. (2005). Enhancing context awareness in augmented reality environments: a review of opportunities and challenges in object and action recognition through machine learning. *Discover Applied Science*, 7, 750. <https://doi.org/10.1007/s42452-025-07033-x>
- [6] Chen, X. (2021). Sustainable development of EFL/ESL learners' willingness to communicate: The effects of teachers and teaching styles. *Sustainability*, 14(1), 396. <https://doi.org/10.3390/su14010396>
- [7] Chun, C. W. (2016). The contribution of critical pedagogy to understanding the politics of English language teaching for sustainable development. *Journal of Applied Linguistics and Professional Practice*, 11(3), 251–272. <https://doi.org/10.1558/japl.v11i3.30161>
- [8] Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- [9] Eriksson, A.-M. (2015). Crafting claims about sustainability assessment: The situated practice of guiding university students into producing and presenting conclusions. *Linguistics and Education*, 30, 97–113. <https://doi.org/10.1016/j.linged.2015.03.011>
- [10] Gahshan, N., & Weintraub, N. (2024). Effect of an Instructional Program for Word Processing and efficient typing on 'Year 4 students' composition. *Reading & Writing*. <https://doi.org/10.1007/s11145-024-10613-9>
- [11] Geyer, J. J. (1972). Comprehensive and Partial Models Related to the Reading Process. *Reading Research Quarterly*, 7(4), 541. <https://doi.org/10.2307/747105>
- [12] Hilterbran, A. (2016). Ana Gimeno, Mike Levy, Françoise Blin, and David Barr (eds): Worldcall: Sustainability and computer-assisted language learning. *Applied Linguistics*, 38(2), 271–275. <https://doi.org/10.1093/applin/amw037>
- [13] Jin, S., & Li, H. (2025). Trends and sustainability of semantic rhetoric translation research: Insights from bibliometric analysis. *Southern African Linguistics and Applied Language Studies*, 1–20. <https://doi.org/10.2989/16073614.2025.2508465>
- [14] Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- [15] Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

- [16] Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size, and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- [17] Lin, Z., & Lei, L. (2020). The research trends of multilingualism in applied linguistics and education (2000–2019): A bibliometric analysis. *Sustainability*, 12(15), 6058. <https://doi.org/10.3390/su12156058>
- [18] Mlinarič, M., Los, S. A., & Snell, J. (2025). On the spatial limits of parallel word processing in reading. *Attention, Perception, and Psychophysics*, 87, 1704–1716. <https://doi.org/10.3758/s13414-025-03101-x>
- [19] Nassaji, H. (2011). The role of vocabulary learning in second language reading. *Language Teaching*, 44(4), 401–423. <https://doi.org/10.1017/S0261444810000202>
- [20] Nervino, E., Cheung, J. O., & Chen, J. (2024). Charting the path of sustainability discourse research: A systematic review of applied linguistic studies. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12537>
- [21] Odia, U. L. (2021). Impact of recycling sustainability on organizational performance. *Linguistics and Culture Review*, 6, 93–105. <https://doi.org/10.21744/lingure.v6ns1.1977>
- [22] Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- [23] Perfetti, C. A., & Stafura, J. (2014). Word knowledge in theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- [24] Piga, A. (2023). Systemic functional linguistics in the rhetorical strategies of persuasion: A longitudinal study of transitivity and ergativity in the rhetoric of Saras' sustainability reports. *International Journal of English Linguistics*, 13(4), 24. <https://doi.org/10.5539/ijel.v13n4p24>
- [25] Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- [26] Rayner, K., Reichle, E. D., & Pollatsek, A. (2006). Eye movements in reading: A tutorial review. *Handbook of Psycholinguistics*, 2, 56–97.
- [27] Sadoughvanini, S., & Ghaemi, H. (2025). Online interaction learning: A way towards sustainability and buoyant EFL students. *Australian Journal of Applied Linguistics*, 8(3), 102494. <https://doi.org/10.29140/ajal.v8n3.102494>
- [28] Schraw, G., & Bruning, R. (1996). Readers' Implicit Models of Reading. *Reading Research Quarterly*, 31(3), 290–305. <https://doi.org/10.1598/rq.31.3.4>
- [29] Snell, J., & Grainger, J. (2018). Parallel word processing in the flanker paradigm has a rightward bias. *Attention, Perception, & Psychophysics*, 80(6), 1512–1519. <https://doi.org/10.3758/s13414-018-1547-2>
- [30] Sterling, S. (2010). Transformative learning and sustainability: Sketching the conceptual ground. *Learning and Teaching in Higher Education*, 5(11), 17–33.
- [31] Tsaprouni, E., & Manouilidou, C. (2025). The Role of Aspect During Deverbal Word Processing in Greek. *J Psycholinguist Research*, 54, 7. <https://doi.org/10.1007/s10936-024-10112-6>
- [32] Wang, L. (2023). Adoption of the PICRAT model to guide the integration of innovative technologies in the teaching of a linguistics course. *Sustainability*, 15(5), 3886. <https://doi.org/10.3390/su15053886>
- [33] Yang, X., & Peng, Z. (2022). Metaphor, sustainability, transformation: Transdisciplinary perspectives. *International Journal of Applied Linguistics*, 33(1), 76–79. <https://doi.org/10.1111/ijal.12437>

Mohammad A. Abdulaal is a faculty member at Prince Sattam bin Abdulaziz University, specializing in linguistics, applied linguistics, and theoretical linguistics. His research has garnered 131 citations, reflecting his contributions to the field. Dr. Abdulaal's research interests include dynamic assessment in EFL learning, psycholinguistic case studies, and cross-linguistic analyses of formulaic language. His work has been published in journals such as *Language Testing in Asia* and the *International Journal of English Linguistics*. A member of the academic community at Prince Sattam bin Abdulaziz University, Dr. Abdulaal contributes to the advancement of linguistic studies through both teaching and research.

Ahmad A. Sharaf Eldin is an Educational Consultant based in the United Arab Emirates, holding a Ph.D. in Linguistics and a specialized degree in Simultaneous Interpretation. His academic journey is marked by significant global engagement in Turkey, Bahrain, and Egypt, alongside prestigious appointments in the United States as a Fulbright Scholar at SUNY and a faculty member at William & Mary, Wake Forest, and Middlebury College. Beyond his work as a certified interpreter for the Egyptian Ministry of Foreign Affairs, Dr. Sharaf Eldin is a recognized media personality who has hosted television programs on linguistics and appeared on various international satellite channels. He is also an accomplished author, having published books and research papers in both Arabic and English that analyze the complex interplay of language, politics, and society.

Hanan M. Al-Johani is an Assistant Professor in the Department of English Language and Literature at the College of Science and Humanities, Prince Sattam bin Abdulaziz University, Saudi Arabia. She earned her Doctor of Education (EdD) degree from the University of Strathclyde in 2009, with a thesis titled "Finding a Way Forward: The Impact of Teachers' Strategies, Beliefs and Knowledge on Teaching English as a Foreign Language in Saudi Arabia." Dr. Al-Johani's research interests include language assessment, virtual learning, and sociolinguistics. She has published works such as "The Impact of the Dynamicity and Non-dynamicity of Assessment on EFL Learners' Productive Skills: Attitude in Focus" and "Social Network Misinformation and Attitudinal Shift: A Sociolinguistic Perspective".

Mohammad A. Aldawood is an Assistant Professor of Linguistics in the Department of English, Faculty of Arts at Port Said

University, Egypt. He holds a PhD in Linguistics with a specialization in pragmatics and discourse studies. His academic interests include pragmatic theory, media discourse, and sociolinguistics in Arab and global contexts. He has authored and co-authored numerous articles in high-impact international journals on language, ideology, and communication. He has participated in several international research collaborations and academic conferences. His work aims to bridge theoretical linguistics with real-world communicative practices in media and public discourse.

Video-Based Platforms and Speaking Skill Development in Undergraduate EFL Contexts: Insights From Flipgrid

Maytham Hashim Mutlaq

Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

Harmi Izzuan bin Baharum

Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

Ali Hamed Barghi

Department of Teaching, Learning and Culture, Texas A&M University, College Station, USA

Seriaznita Mat Said

Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia

Abstract—Many EFL learners do not have enough opportunities to practice speaking English in real situations. This study examined how Flipgrid, a video discussion platform, influenced the speaking performance and perceptions of undergraduate EFL learners. Sixty Iraqi university students from an English department participated in an eight-week mixed-methods study. The WIDA Speaking Rubric was used to assess students in pre and post speaking tests, which provided quantitative data, while semi-structured interviews with ten participants provided the qualitative data. The study showed that students improved their speaking skills, especially in fluency and pronunciation. The analysis also showed progress in vocabulary, grammar, and comprehension. During the interviews, many students said that using Flipgrid made them feel more confident and more willing to speak English. A few also mentioned that they felt less nervous when speaking. Students appreciated recording videos more than once, practicing independently, and getting feedback from both peers and teachers. Peer feedback allowed students to notice language patterns and think about different ways to express their ideas. Some participants pointed out issues such as unstable internet connections, limited experience with technology, and anxiety about recording themselves. Cultural factors also made it difficult to include video tasks in regular lessons.

Index Terms—speaking skills, video-based platforms, EFL learners, educational technology, Flipgrid

I. INTRODUCTION

Teachers today use many kinds of digital tools to make language learning more engaging. In EFL classrooms, video platforms help students interact and create a space where they can speak more freely. Flipgrid was popular until September 2024 because it was easy to handle. Students could record and share short video responses with almost no technical effort. After that date its core functions moved into Microsoft Teams for Education through Microsoft Reflect and Video Journal. These tools include transcription and translation. Even so many teachers describe them as less intuitive and less convenient for everyday classroom work. This study examines how Flipgrid helped students in Iraqi EFL classrooms improve their speaking skills. It describes the ways the platform supported language learning and offers insights that may guide teachers in using similar tools more effectively. The findings aim to promote learner participation, build confidence, and enhance speaking ability.

Using technology in EFL classrooms can help with challenges such as limited speaking practice, student anxiety, and low participation (Ahmed, 2019; Al-Abdali & Alzayadi, 2020; Lamb & King, 2020). Among these tools, video-based platforms are especially effective in promoting interaction and learner-centered activities (Lamb & King, 2020). Studies also show that the quality of teacher support and the satisfaction of learners' needs for competence, autonomy, and relatedness strongly influence motivation and achievement in online environments (Ellis, 2009; Richards & Rodgers, 2001). Such findings confirm that the pedagogical design surrounding technology is as important as the tool itself.

Some learning theories explain why such platforms are effective. Constructivist theory suggests that students learn more effectively when they engage in meaningful activities, interact with peers and teachers, and reflect on their learning (Denzin, 1978; WIDA, 2020). Yet many EFL teachers, though familiar with learner-centered approaches, find it difficult to apply them in practice because of local constraints and fixed curricula (Kaivanpanah et al., 2019). Flipgrid provides structured spaces that help learners reflect and work more independently. Research also indicates that multimedia tasks

that encourage reflection and independent learning can improve pronunciation and help students become more self-directed (Ghafar & Amin, 2022).

According to the Unified Theory of Acceptance and Use of Technology (UTAUT), students are more willing to use educational technologies when they see them as useful, easy to handle, and supported by teachers and peers (Venkatesh et al., 2012; Ahmed et al., 2017). Flipgrid's simple setup lets students record short videos, watch their classmates' recordings, and give feedback in written or video form. These functions encourage active participation and consistent practice (Ghareb & Mohammed, 2017; Naeem & Saad, 2020).

Task-Based Learning (TBL) highlights learning through activities that represent real communication and authentic language use (Phuong, 2022). In this study, Flipgrid was applied for weekly speaking tasks connected to students' everyday experiences and cultural contexts. Learners recorded short videos on topics such as family traditions, used sentence starters to guide their speech, and then provided feedback on their classmates' work. Through these exchanges, students became more fluent, more accurate, and more confident in speaking (Lamb & King, 2020; Phuong, 2022).

Previous studies have also highlighted the value of video-based platforms. Lamb and King (2020) found that such tools improved students' fluency, pronunciation, and confidence. Stoszkowski et al. (2021) found that using video-based activities helped students work together more easily and feel closer as a class, which made the learning process more effective. Along with previous research on online need satisfaction (Ellis, 2009; Richards & Rodgers, 2001; Stoszkowski et al., 2021), these findings show that technology is most effective when it nurtures learners' motivation and sense of belonging.

In Iraq, student participation is often limited by teacher-centered instruction, large class sizes, and strict curricula (Ghafar & Amin, 2022). In many cases, gender expectations make students less comfortable speaking when classes include both males and females (Ahmed et al., 2017). Limited internet service and few available devices often prevent students from using online learning tools effectively (Ghareb & Mohammed, 2017). Even so, research shows that technology can help students take part more actively and improve their speaking skills (Naeem & Saad, 2020; Mussa, 2020). Despite these findings, evidence from Iraqi EFL classrooms remains limited, and little is known about how undergraduate students experience video-based speaking tasks in such challenging conditions.

In most Iraqi EFL classrooms, lessons are directed mainly by the teacher, and students have limited chances to review or respond to each other's work. Introducing video tools changes this situation, as students can now view classmates' work, exchange comments, and learn from one another in ways that ordinary classroom settings rarely make possible. This change encourages students to take a more active role in their learning rather than depending solely on the teacher. Studies show that using digital tools for peer feedback can make classrooms feel less hierarchical (Naeem & Saad, 2020; Mussa, 2020). When the focus shifts this way, students often feel more at ease trying new language and taking risks (Ahmed, 2019; Al-Abdali & Alzayadi, 2020). Still, the effectiveness of this kind of interaction depends on whether students know how to give useful feedback and whether the class atmosphere encourages cooperation. When peer review is new to a class, teachers may need to model it first and guide students step by step until they can manage the process on their own.

Taking these challenges and opportunities into account, this study looks at how Flipgrid supports the development of Iraqi undergraduate EFL students' speaking skills. It focuses on answering three main research questions:

1. How do students perceive the use of Flipgrid in improving their speaking skills?
2. How does Flipgrid influence speaking proficiency in terms of fluency, pronunciation, vocabulary, grammatical accuracy, and comprehension?
3. What challenges emerge during the integration of video-based platforms in EFL classrooms?

Having outlined the theoretical basis for using video-based platforms, the next section describes the research methods and data collection procedures.

II. METHOD

Research Design

This study used a sequential transformative mixed methods design to investigate how Flipgrid supported speaking skill development among Iraqi EFL learners. The design combined quantitative and qualitative approaches, where quantitative data provided an initial measure of progress, and qualitative data explained learners' experiences in depth (Creswell, 2014). This approach was chosen because it helps the researcher examine both measurable results and participants' real experiences, which is consistent with recent recommendations in EFL technology integration research (Phuong, 2022; Ghafar & Amin, 2022).

Participants

The participants were 60 undergraduate EFL students (40 male and 20 female) aged 19 to 21 from the University of Baghdad. They were selected from an intact class using convenience sampling to represent real classroom conditions. All had studied English since primary school and showed mixed proficiency levels typical of Iraqi EFL contexts. Participation was voluntary, and informed consent was obtained prior to data collection. Ethical approval was granted by the university's English Department Research Committee.

Researcher's Role

The researchers acted as both facilitators and observers during the study. One of the researchers, with more than ten years of experience teaching EFL, conducted the Flipgrid workshops and followed the students' progress each week. Familiarity with the students' academic culture helped build trust and encouraged participation but also posed a risk of interpretive bias. To minimize this, both researchers kept reflective memos and reviewed their analyses collaboratively (Lincoln & Guba, 1985), which helped ensure balanced interpretation and credibility of the findings.

Instruments

Two main instruments were used: (a) the WIDA Speaking Rubric (Lowenthal & Moore, 2020) for pre- and post-assessment, and (b) a 20-item perception questionnaire. The WIDA rubric evaluates five areas: fluency, pronunciation, vocabulary, grammar, and task performance, using a six-point scale. Scores were assigned independently by two raters trained in EFL assessment, and inter-rater reliability was calculated at 0.92 (Cohen's Kappa), showing high agreement. Disagreements were discussed until consensus was reached (Landis & Koch, 1977; WIDA, 2020). The perception questionnaire measured learners' attitudes toward Flipgrid using a five-point Likert scale. Cronbach's alpha reliability reached 0.96, indicating strong internal consistency (George & Mallery, 2003).

Data Collection and Procedures

The data were collected in three phases, which took place before, during, and after the Flipgrid intervention. The quantitative phase, which included pre- and post-tests, was completed first to provide a measurable baseline and to guide the later qualitative phase.

In the first week, students took part in three 90-minute workshops held in the computer lab, where they practiced recording, uploading, and responding to videos. Through guided activities and demonstrations, all participants developed the basic technical skills needed before starting the main program. Throughout the intervention, the researcher also observed classroom sessions and recorded field notes to capture learners' participation patterns and engagement.

Students' speaking proficiency was assessed at the beginning and end of the eight-week program using the WIDA rubric (Lowenthal & Moore, 2020). All recording sessions were carried out under the same classroom conditions to ensure fair evaluation.

The Flipgrid activities were designed based on a task-based learning framework (Ellis, 2009; Richards & Rodgers, 2001). Each week, students created a short video on everyday topics such as family traditions or cultural customs and then received feedback from both peers and the teacher. The CEFR descriptors ensured tasks matched learners' proficiency levels, and scaffolding (sentence stems and visuals) was gradually withdrawn to encourage autonomy (Council of Europe, 2001).

After the intervention, a post-test was given, and students were classified based on their level of improvement as high, medium, or low. From these groups, ten students were intentionally chosen for semi-structured interviews, including four from the high group, three from the medium group, and three from the low group. Altogether, ten interviews were carried out, each lasting about twenty to twenty-five minutes, producing nearly two hundred minutes of audio data.

Data Analysis

Quantitative data from the pre- and post-tests were analyzed using paired sample t-tests to measure improvement in fluency, pronunciation, and vocabulary. Questionnaire reliability was confirmed through Cronbach's alpha ($\alpha = 0.96$).

Qualitative data were analyzed following Braun and Clarke's six-phase thematic framework (Braun & Clarke, 2006). The transcripts were read repeatedly to identify key ideas, which were assigned descriptive codes and later grouped into broader categories such as confidence building and peer collaboration. These categories were refined into final themes that represented students' shared experiences. Two coders completed the analysis independently and compared their codes to ensure consistency, reaching a Cohen's Kappa of 0.88 (Landis & Koch, 1977). Triangulation of test scores, questionnaire results, interviews, and researcher observations further strengthened the validity of the findings (Denzin, 1978; Lincoln & Guba, 1985).

Validity and Reliability

To ensure methodological rigor, multiple procedures were applied. Inter-rater reliability confirmed that raters assigned consistent scores for speaking performance. To reduce bias, the researchers linked quantitative data, qualitative findings, and personal observations through triangulation. Member checking was conducted by summarizing the main interview themes and inviting participants to confirm or clarify their views. Their confirmations helped verify the accuracy of interpretations. These procedures are consistent with current recommendations for enhancing credibility in EFL mixed-methods research (Creswell, 2014; Denzin, 1978; Lincoln & Guba, 1985).

III. RESULTS

This section presents the study's findings, combining statistical analyses with qualitative insights. Each research question is addressed separately.

How do undergraduate EFL students perceive the use of Flipgrid in improving their speaking skills?

Students' perceptions of Flipgrid were evaluated using a 20-item questionnaire. The questionnaire used a five-point Likert scale in which participants chose responses ranging from 1, meaning strongly disagree, to 5, meaning strongly agree.

The reliability of the questionnaire was checked by examining how consistently the items measured the same construct. The overall Cronbach's alpha was 0.96, and item alphas ranged from 0.81 to 0.96, showing strong reliability (George & Mallery, 2003). These findings are shown in Table 1.

TABLE 1
RELIABILITY ANALYSIS OF THE FLIPGRID PERCEPTION QUESTIONNAIRE

Components	Cronbach's Alpha
Overall	0.96

Table 2 lists the questionnaire items and their results, showing the mean scores and standard deviations for each one.

TABLE 2
MEAN SCORES AND STANDARD DEVIATIONS FOR PARTICIPANTS' PERCEPTIONS OF FLIPGRID'S IMPACT ON SPEAKING SKILLS

S/N	Items	M	SD
1	I believe that practicing English speaking using Flipgrid has helped increase my cooperation and communication with my classmates.	4.01	.33
2	I believe that Flipgrid makes me autonomous in English-speaking practices.	3.99	.35
3	I believe that learning English speaking will be effective with the use of Flipgrid.	4.04	.53
4	I am less frightened about making mistakes when learning English speaking using Flipgrid.	3.81	.63
5	I consider Flipgrid a great English learning tool.	3.89	.53
6	I feel comfortable practicing speaking English using Flipgrid.	3.95	.60
7	I believe that Flipgrid has helped me reduce my nervousness in learning English speaking.	3.87	.66
8	I believe that the use of Flipgrid has helped me become self-confident in my speaking performance.	3.90	.59
9	I am responsible for my English learning when using Flipgrid.	3.99	.57
10	I believe that Flipgrid has helped me speak English more fluently.	3.96	.66
11	I believe that Flipgrid has helped me better improve my pronunciation.	3.88	.58
12	I believe that practicing English speaking using Flipgrid has helped me better communicate with my teachers.	3.95	.59
13	I believe that practicing English speaking using Flipgrid has helped me recognize mistakes.	3.92	.56
14	Learning using Flipgrid encouraged me to practice speaking in English.	3.95	.71
15	I would like to study English speaking without the use of Flipgrid.	3.81	.63
16	I hope Flipgrid is used more frequently to make English-speaking learning more fun.	4.01	.33
17	I hope Flipgrid will enable me to study English speaking more efficiently.	3.89	.53
18	I think I will continue using Flipgrid for English-speaking learning in the future.	3.99	.35
19	I find Flipgrid more engaging than traditional speaking practice.	3.85	.62
20	I believe Flipgrid helps me track my speaking progress over time.	3.90	.58

Students generally viewed Flipgrid as a helpful way to improve their speaking. Mean scores of 4.04 for effectiveness and 4.02 for enjoyment show that most students were motivated and pleased with the video speaking activities (Nagy, 2018).

Interview data supported these quantitative results. Student S01 mentioned, "I liked that I could redo my recordings because it made me less worried about making mistakes." Student S07 shared, "Giving feedback on my classmates' videos made speaking practice more enjoyable." These comments are consistent with findings from recent studies suggesting that learner-centered digital tools can lower performance anxiety and build a stronger sense of autonomy (Phuong, 2022, 2023, 2025).

Some students also noted difficulties. Student S03 noted, "It was hard to connect the Flipgrid tasks with what we were doing in class," which points to the need for stronger institutional support and clearer guidelines for integrating such tools (Ghareb & Mohammed, 2017).

Most students said that using Flipgrid made them feel more confident and less afraid to speak. Their average scores for lower anxiety ($M = 3.88$, $SD = 0.67$) and higher confidence ($M = 3.91$, $SD = 0.60$) support this feeling.

Flipgrid allowed students to practice speaking at their own pace without feeling pressured. The interviews highlighted three key themes. Students worked together more effectively, gained confidence, and described the experience as enjoyable. They liked having the chance to record their responses again, practice independently, and receive feedback from peers and the teacher in a comfortable environment. The simple layout of the platform also made the learning process smoother. Overall, the statistical results and interview findings showed that Flipgrid increased students' motivation, participation, and confidence, confirming its value as a useful tool for improving speaking skills in EFL classrooms (Phuong, 2022, 2023, 2025).

How does Flipgrid influence speaking proficiency in terms of fluency, pronunciation, vocabulary, grammatical accuracy and comprehension?

Speaking proficiency was evaluated through pre- and post-tests that used the WIDA Speaking Rubric (WIDA, 2020). Scores were given for fluency, pronunciation, vocabulary, grammar, and comprehension. A paired-sample t-test showed significant improvement in all five dimensions ($p < .001$).

Two trained raters independently evaluated all recordings. Inter-rater agreement was strong, with ICC = 0.873 (95 % CI = 0.760-0.951, $F = 8.544$, $p < .001$), which exceeds the threshold for substantial agreement (Landis & Koch, 1977). Table 3 gives a detailed summary of these findings.

TABLE 3
INTER-RATER RELIABILITY ANALYSIS FOR SPEAKING SKILL ASSESSMENT

Measure	ICC	95% Confidence Interval	F	df1	df2	p
Average Measures	.873	.760 - .951	8.544	17	51	.000

The comparison of means showed gains in all areas (Table 4). Fluency showed the highest improvement ($M = 1.56$, $p < .001$, $d = 2.24$). Pronunciation also improved well ($M = 1.10$, $p < .001$, $d = 3.64$). Gains were smaller but still clear in vocabulary ($M = 0.67$, $p < .001$, $d = 0.98$), grammar ($M = 0.93$, $p < .001$, $d = 1.08$), and comprehension ($M = 0.72$, $p < .001$, $d = 0.91$).

TABLE 4
PRE- AND POST-TEST RESULTS FOR SPEAKING SKILL IMPROVEMENTS (PAIRED-SAMPLE T-TEST)

Components	Pretest <i>M</i>	Post-test <i>M</i>	Mean Diff.	<i>SD</i>	<i>t</i>	<i>P</i> (2 tailed)	Effect Size (Cohen's <i>d</i>)
Fluency	2.93	4.50	1.56	.69	-17.38	.000	2.24
Pronunciation	2.00	3.10	1.10	.302	-28.16	.000	3.64
Vocabulary	2.93	3.60	.666	.680	-7.587	.000	0.98
Grammatical Accuracy	2.86	3.80	.933	.860	-8.397	.000	1.08
Comprehension	3.00	3.71	.716	.783	-7.08	.000	0.91
Overall	13.73	18.71	4.98	1.92	-20.03	.000	1.77

($n=60$)

Qualitative data from the interview explained these patterns. Students explained that making several recordings helped them speak with greater ease. One student said that doing the task several times made speaking feel smoother and more natural. Another student (S05) explained that feedback from the teacher played an important role in improving pronunciation. Prompts and word lists made it easier to use new vocabulary, and teacher guidance helped students focus more on grammar.

Figure 1 shows the pre-test and post-test averages for the five areas of speaking skill. The findings show clear progress, especially in fluency and pronunciation.

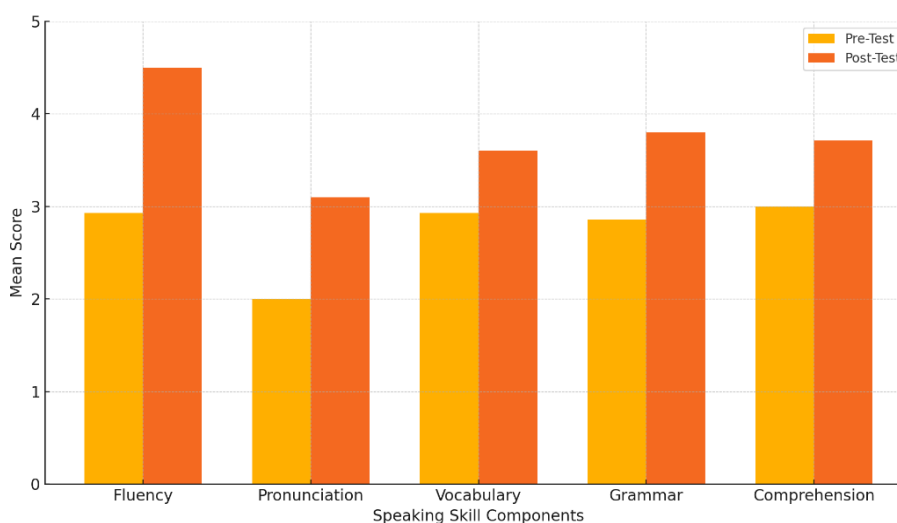


Figure 1. Pre- and Post-Test Means Across Components of Speaking Proficiency

In sum, both statistical and qualitative evidence indicate that regular Flipgrid practice, constructive feedback, and reflective recording fostered measurable progress in students' oral proficiency (WIDA, 2020).

What challenges emerge during the integration of video-based platforms in EFL classrooms?

Flipgrid supported learning, but it also presented several challenges. Teachers and students faced technical issues, weak institutional support, and cultural barriers that made it less effective. One main issue was that the platform could do more

than the institution was ready to support. Many teachers found it difficult to integrate Flipgrid activities into their lessons due to limited guidance and insufficient resources. Student S11 mentioned that making Flipgrid tasks align with classroom lessons was challenging. Another student mentioned that the school gave little direction on how to use Flipgrid in teaching.

Students' experiences were not the same. Students who often used applications like WhatsApp or Instagram learned to use Flipgrid more easily. Others who had limited experience with online tools needed extra time and support to get used to it. Weak internet connections also made participation harder. Students reported slow internet and weak devices; for instance, S08 said, "My videos took too long to upload, and I missed deadlines".

Cultural factors presented additional barriers. Some female participants felt uneasy about recording videos that male classmates could view, consistent with findings on gendered participation constraints in Iraqi classrooms (Ahmed et al., 2017; Ghafar & Amin, 2022). Repetition of tasks and the relevance of topics also affected students' motivation. They were more interested when the tasks related to their own lives but lost focus when the topics were unfamiliar or repeated. In general, Flipgrid improved students' speaking ability and confidence, yet its continued success relies on teacher preparation, institutional backing, and sensitivity to cultural factors (Creswell, 2014; Ghareb & Mohammed, 2017). The next section discusses these findings in relation to earlier research and considers their implications and limitations.

IV. DISCUSSION

The first research question looked at how undergraduate EFL students perceived Flipgrid and its effect on their speaking skills. The findings indicated that students considered Flipgrid both useful and encouraging. Students appreciated being able to record at their own pace, practice several times, and reflect on their speaking without the pressure of immediate evaluation. Feedback from classmates also enriched the experience, giving them the opportunity to exchange ideas and learn in a supportive environment. These results align with constructivist views that emphasize learning through active participation and interaction (Piaget, 1970), as students built understanding through ongoing reflection and social engagement. These findings are in line with previous research (Stoszkowski et al., 2021), which found that digital learning environments support learner independence and reflection when students are given enough time to review their work.

Students felt less anxious and more confident, which highlights the social side of language learning where interaction and supportive feedback reduce the fear of making mistakes. This result is consistent with studies on technology-based speaking platforms that build learners' confidence through repeated practice and self-assessment opportunities (Lamb & King, 2020; Lowenthal & Moore, 2020; Richards & Rodgers, 2001; Stoszkowski et al., 2021). Phuong (2025) also found that taking part regularly in online reflective activities increases learners' sense of responsibility, which helps them become more confident and persistent in communication. A more recent study by Hwang et al. (2025) also found that students working with video-based speaking tasks were more actively involved in lessons and showed greater motivation to participate than peers in conventional speaking classes. Observations of Hwang et al. (2025) suggest that when speaking tasks are built around video and repetition, learners are more willing to stay engaged and interact confidently over time. In this study, students' willingness to participate more actively reflects the principles of technology acceptance theory (Ahmed et al., 2017; Venkatesh et al., 2012), where positive experience and perceived usefulness encourage continued engagement.

The peer feedback feature of Flipgrid played an important role because it encouraged a kind of collaborative learning that was different from regular classroom interaction. Students practiced speaking but also spent time listening and responding to one another. When students acted in both roles, they began to notice language features they usually missed when focusing only on their own speech. As they viewed classmates' videos, they learned other ways to phrase ideas and became more aware of how language choice changes what a message sounds like. In the interviews, a few students said they felt more comfortable getting feedback from their peers than from the teacher. This pattern shows that video-based tasks help make learning something students do together instead of something led only by the teacher. They also helped students take more part in shaping class activities, which is important where lessons are often teacher-centered. These results agree with the task-based learning view that real communication tasks help learners share meaning and notice language form through use (Ellis, 2009; Richards & Rodgers, 2001). They also support the constructivist view that learning develops through working with others.

Since the study participants came from a single Iraqi university, generalization must be done carefully. Contextual factors such as institutional readiness, access to technology, and cultural expectations may influence adoption and performance. Participants' initial hesitation and varying degrees of digital familiarity show that technology alone cannot guarantee learning gains. Guidance and continuous training are necessary for students to make full use of digital tools, as noted in previous Iraqi research (Piaget, 1970; Venkatesh et al., 2012). This finding also aligns with the technology acceptance perspective (Ahmed et al., 2017; Venkatesh et al., 2012), which emphasizes that users' confidence, perceived usefulness, and support conditions affect their engagement with new tools. In such contexts, teachers can include basic digital literacy activities in English courses so that students learn how to use online tools with more ease.

The second research question looked at how Flipgrid affected students' fluency, pronunciation, vocabulary, grammar, and comprehension. Evidence from both quantitative results and interviews showed progress in all five areas, with the greatest improvement in fluency. Students benefited from being able to record several times and review their work, which helped them speak more smoothly without time pressure. These results support the idea that regular communication practice in real situations improves fluency (Ellis, 2009; Jamil et al., 2025; Richards & Rodgers, 2001; Stoszkowski et al.,

2021). The improvement in pronunciation also shows that listening to their own and others' recordings helped students notice and correct their pronunciation, similar to what Lamb and King (2020) described in their studies on collaborative learning. This pattern follows task-based learning principles, in which repeated practice and reflection help students build fluency and accuracy through real language use (Ellis, 2009; Richards & Rodgers, 2001). A comparable outcome was noted in a recent university-based study; Al-Hamed et al. (2025), in working with EFL learners, found that regular use of video conversation activities was associated with noticeable gains in fluency and pronunciation. They also found that allowing students to rehearse, listen to their own recordings, and speak without immediate classroom pressure played an important role in shaping EFL learners fluency and pronunciation.

In this study, fluency and pronunciation improved more than other areas, while progress in vocabulary and grammar was smaller. This shows that frequent practice and communication can help students remember words and use grammar better, but they still need lessons that focus on form for long-term accuracy. Ellis (2009) noted that meaning-based activities help learners speak more easily and interact better, whereas form-based instruction helps them become more accurate. These results support that view and agree with Ahmed (2019), who highlighted the importance of combining communication and grammar-focused teaching in EFL classrooms. In our study, the mix of communicative video tasks and reflective feedback created conditions where meaning and form complemented each other, which is consistent with task-based learning principles (Ellis, 2009; Richards & Rodgers, 2001).

The results further confirm previous findings (Ahmed et al., 2017; Ghareb & Mohammed, 2017; Naeem & Saad, 2020) that digital platforms are most effective when embedded in guided learning frameworks. Students' progress came not only from using Flipgrid but also from the pedagogical design that joined technology with reflective, scaffolded, and socially supportive activities. Planned use of prompts and teacher feedback guided students to focus on both language form and meaning at the same time, which supports task-based learning theory (Ellis, 2009; Richards & Rodgers, 2001). These findings align with previous studies showing that organized technological guidance helps learners become more independent and stay engaged (Ahmed et al., 2017; Ghareb & Mohammed, 2017; Naeem & Saad, 2020). In this way, Flipgrid extended classroom learning by helping students work both on their own and with others.

The study found that the design of a task influences how students engage with it. Many students said they felt more motivated when activities related to their own lives, for instance, when they spoke about family customs or everyday traditions. This shows that video-based learning works best when teachers connect the tasks to students' real experiences rather than just the tool itself. When students see what they speak about is related to their own lives, they use the language more thoughtfully. This finding aligns with constructivist and sociocultural perspectives (Denzin, 1978; WIDA, 2020), which stress that learning becomes meaningful when new knowledge links to learners' personal and cultural experiences. It also supports earlier research showing that lessons built around students' cultural backgrounds can make them more active and improve learning. Teachers in similar classrooms could prepare topics that reflect what students know and care about, so speaking online feels closer to real communication than a school task.

In using Flipgrid during lessons, teachers and students mentioned that slow internet and limited access to devices made participation difficult. This finding is consistent with Ghareb and Mohammed's (2017) report that weak infrastructure continues to restrict the effective use of technology in Iraqi universities. These challenges also underline the importance of supporting conditions in the technology acceptance framework (Ahmed et al., 2017; Venkatesh et al., 2012), which indicates that access to resources and technical help shapes how ready users are to adopt digital tools. Universities can respond to these challenges by improving internet access, providing the necessary equipment, and guiding both teachers and students in using technology effectively.

Teachers also spoke about challenges they faced with instruction. Many reported difficulty incorporating Flipgrid activities into their lessons due to insufficient support and guidance from administrators. This agrees with Lamb and King's (2020) idea that technology works best in education when institutions plan clearly, prepare teachers well, and connect digital tools with curriculum aims. Phuong (2025) also found that when schools do not plan carefully, teachers tend to return to older methods of instruction, which reduces the benefits of digital learning. These findings suggest that institutional support is essential for sustaining task-based and reflective teaching practices, as successful technology integration depends on both pedagogical design and systemic readiness.

One factor that influenced how effectively students used Flipgrid was their varied levels of digital literacy. Some students who were used to social media picked it up quickly, while others needed more time and help. This difference shows why teachers must think carefully about how they use technology with learners from mixed backgrounds. Introducing a new tool does not ensure that every student will benefit equally. In the beginning, teachers may need to spend more time helping those who are less familiar with technology learn how to record, upload, and respond to videos. Giving simple demonstrations or pairing tech-confident students with peers who need support can make the process smoother and more inclusive. When teachers notice these gaps early and take steps to address them, all students can benefit more equally from video-based speaking activities, including those who are less familiar with technology. This peer-assisted approach also reflects constructivist and social learning views, where collaboration and shared problem-solving enhance understanding (Denzin, 1978; WIDA, 2020).

The study also shows that both teachers and students need preparation before using video tools in language classes. Technical training alone does not make the process work well. Teachers need practical help with planning useful tasks, giving feedback, and grading video work fairly. Students need to learn more than how to use the platform. They should

also know how to give thoughtful comments on each other's work and reflect on their own progress. Schools that take time to build these habits early tend to see better results. Without this kind of preparation, the platform may end up as a tool that is rarely used instead of something that truly supports learning. These insights are consistent with reflective and transformative learning perspectives (Denzin, 1978; Phuong, 2022), which highlight how guided reflection and deliberate engagement can transform digital activities into opportunities for meaningful professional and personal development. Future programs that introduce technology into EFL classrooms should provide training that covers both technical skills and effective teaching practices in digital learning.

Dealing with these challenges calls for a coordinated effort. Schools need to train teachers in task design, digital assessment, and reflective teaching, and they should also create policies that explain clearly how technology fits into the learning process. Moreover, students need access to continual technical support to ensure equitable participation. These measures would strengthen both confidence and competence, allowing digital learning platforms like Flipgrid to reach their full potential. Teacher training programs should move beyond introducing technical features and focus instead on helping educators understand how to integrate video-based activities into lesson plans in ways that support specific learning objectives. This means showing teachers how to design tasks that encourage meaningful communication rather than simple performance, how to give feedback that guides improvement without discouraging learners, and how to assess video responses fairly and consistently. At the institutional level, clear policies are needed to handle practical issues such as giving teachers time to review student videos, offering alternative internet access for those with weak connections, and setting rules that protect privacy and cultural values. When schools treat technology integration as a long-term commitment instead of a temporary experiment, they make it easier for teachers and students to develop the skills and confidence needed to use digital learning tools effectively and consistently. This approach reflects a transformative view of educational change, in which sustainable innovation depends on the alignment of teacher preparation, institutional support, and reflective practice (Phuong, 2022).

More broadly, this study adds to the understanding of how video-based platforms can support both language growth and emotional development in EFL learning. By combining constructivist, task-based, and technology acceptance frameworks, the findings show that technology can empower learners to take responsibility for their progress and create meaningful communicative experiences. At the same time, they reveal how sociocultural and infrastructural factors shape this process. These results extend the insights of Phuong (2022, 2023, 2025) by showing that context-sensitive implementation, strengthened through reflection, feedback, and institutional readiness, determines the success of technology-enhanced speaking pedagogy.

Overall, this study shows that video-based platforms such as Flipgrid can help lower students' speaking anxiety, encourage cooperation among learners, and improve speaking skills, although some contextual challenges still remain. The findings reinforce the combined value of constructivist, task-based, and technology acceptance perspectives in explaining how digital tools shape learning outcomes. By focusing on Iraqi EFL classrooms, the study contributes important empirical evidence to an underrepresented context in technology-enhanced language learning research (Ahmed et al., 2017; Creswell, 2014; Ellis, 2009; Ghareb & Mohammed, 2017; Naeem & Saad, 2020; Richards & Rodgers, 2001; Stoszkowski et al., 2021).

V. CONCLUSION

This study provides insight into how video-based platforms can help Iraqi undergraduate EFL students improve their speaking. The findings indicate that such tools help build a supportive environment that develops fluency, pronunciation, confidence, and learner independence. These outcomes correspond with constructivism, task-based approach to language learning, and technology acceptance perspectives, which explain how digital platforms can foster meaningful communication and reflective learning. The study also identified several local challenges, such as technical difficulties, cultural considerations, and instructional constraints, which should be addressed to support the effective and sustainable use of these tools.

Flipgrid stopped operating in September 2024, and its main features were transferred to Microsoft Teams for Education. These findings are therefore relevant for teachers who use comparable platforms. The study highlights that the effective use of such tools depends on institutional support, cultural sensitivity, and the thoughtful integration of video-speaking tasks with course objectives. The insights drawn from this research go beyond a single platform, as the observed challenges and benefits represent broader trends in the use of video-based speaking tools rather than Flipgrid alone. Teachers who adopt similar technologies should consider how these tools support learners with varying levels of digital competence and how they can be adapted to suit local classroom realities. Institutions that seek to integrate video-based platforms successfully must evaluate both infrastructure readiness and staff training, ensuring that new technologies fit existing educational practices (Tuyet & Khang, 2020; Venkatesh et al., 2012). In this sense, the study supports reflective and transformative approaches to teacher education, showing that meaningful technology integration requires ongoing reflection, planning, and institutional commitment.

Future research could examine whether platforms like Microsoft Teams or other AI supported tools help students improve their speaking skills as effectively as current options, or perhaps even more so. Long term studies could show how learners' language abilities develop in actual classroom settings. It is also important to examine how cultural and institutional factors affect student participation and learning in video based speaking activities, particularly in settings

with fewer resources. More studies are needed to understand how these tools help students develop their vocabulary and grammar when teachers offer clear structure and opportunities for reflection. Conducting such research across different EFL contexts could provide deeper insights into how constructivist, task based, and technology acceptance frameworks work together to explain the educational value of digital platforms (Ahmed et al., 2017; Creswell, 2014; Ellis, 2009; Ghareb & Mohammed, 2017; Naeem & Saad, 2020; Richards & Rodgers, 2001; Stoszkowski et al., 2021).

REFERENCES

- [1] Ahmed, H. R. (2019). The perceptions and attitudes of EFL learners in Kurdistan Region of Iraq towards online English courses. *Research Journal of English Language and Literature*, 7(4), 288–300.
- [2] Ahmed, N., Pathan, Z. H., & Khan, F. S. (2017). Exploring the causes of English language speaking anxiety among postgraduate students of University of Balochistan, Pakistan. *International Journal of English Linguistics*, 7(2), 99–105.
- [3] Al-Abdali, A. I. E., & Alzayadi, A. D. N. (2020). The electronic approaches: The era of emergence e-learning in the educational system in Iraq. *International Journal of English Literature and Social Sciences*, 5(5), 1424–1426.
- [4] Al-Hamed, A. M. J. (2025). The effectiveness of incorporating video conversations in an EFL university classroom for promoting oral communication skills. *International Journal for Scientific Research*, 4(11), 195–231.
- [5] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- [6] Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- [7] Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage.
- [8] Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). McGraw-Hill.
- [9] Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- [10] George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Allyn and Bacon.
- [11] Ghafar, Z. N., & Amin, M. Y. M. (2022). Difficulties in speaking English among EFL students in the Iraqi Kurdistan Region. *World Journal of English Language*, 12(6), 286–293.
- [12] Ghareb, M. I., & Mohammed, S. A. (2017). The future of technology-based classrooms. *UHD Journal of Science and Technology*, 1(1), 27–32.
- [13] Hwang, G. J., Fathi, J., & Rahimi, M. (2025). Fostering EFL Learners' Speaking Skills and Flow Experience with Video-Dubbing Tasks: A Flow Theory Perspective. *Journal of Computer Assisted Learning*, 41(2), e13120.
- [14] Jamil, N. Z., Barghi, A. H., Nimehchisalem, V., & Ali, A. M. (2025). Professional development gaps and needs in speaking assessment for Malaysian English language teachers. *Open Journal of Modern Linguistics*, 15(1), 37–48.
- [15] Kaivanpanah, S., Alavi, S. M., & Barghi, A. H. (2019). Teachers' practices and awareness of grammar teaching principles. *Teaching English Language*, 13(2), 213–246.
- [16] Lamb, P., & King, G. (2020). Another platform and a changed context: Student experiences of developing spontaneous speaking in French through physical education. *European Physical Education Review*, 26(2), 515–534.
- [17] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- [18] Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- [19] Lowenthal, P. R., & Moore, R. L. (2020). Exploring student perceptions of Flipgrid in online courses. *Online Learning*, 24(4), 28–41.
- [20] Mussa, I. H. (2020). An investigation of e-learning motivation of Iraqi secondary students. *International Journal of Education, Humanities and Social Science*, 3(2), 25–36.
- [21] Nagy, J. T. (2018). Evaluation of online video usage and learning satisfaction: An extension of the technology acceptance model. *International Review of Research in Open and Distributed Learning*, 19(1), 160–185.
- [22] Naeem, N., & Saad, N. (2020). Iraqi EFL students' perspective on speaking anxiety: A pilot study. *Asian Journal of Humanities, Arts and Literature*, 7(2), 113–122.
- [23] Phuong, T. T. H. (2022). Learning motivation of Vietnamese EFL learners in fully online classes during COVID-19: A mixed-methods study. *Issues in Educational Research*, 32(2), 681–700.
- [24] Phuong, T. T. H. (2023). Prolonged online learning: An exploratory mixed-methods study on EFL learners' needs and need satisfaction. *International Journal of Emerging Technologies in Learning*, 18(4), 186–201.
- [25] Phuong, T. T. H. (2025). Integration of ICT into EFL classrooms: Have Vietnamese primary schools progressed post-COVID-19? *Issues in Educational Research*, 35(2), 759–788.
- [26] Piaget, J. (1970). *Science of education and the psychology of the child*. Viking Press.
- [27] Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge University Press.
- [28] Stoszkowski, J., Hodgkinson, A., & Collins, D. (2021). Using Flipgrid to improve reflection: A collaborative online approach to coach development. *Physical Education and Sport Pedagogy*, 26(2), 167–178.
- [29] Tuyet, T. T. B., & Khang, N. D. (2020). The influences of the Flipgrid app on Vietnamese EFL high school learners' speaking anxiety. *European Journal of Foreign Language Teaching*, 5(1), 128–149.
- [30] Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178.
- [31] WIDA. (2020). *WIDA speaking rubric Grades 1–12*. Wisconsin Center for Education Research. Retrieved April 4, 2025, from <https://wida.wisc.edu/sites/default/files/resource/WIDA-MODEL-Speaking-Rubric.pdf>

Maytham Hashim Mutlaq is a PhD candidate at Universiti Teknologi Malaysia and currently serves as an English Lecturer in Applied Linguistics at Middle Technical University in Iraq. His current research focuses on the role of technology in English language teaching and learning, with particular attention to improving the speaking skills of EFL learners.

Harmi Izzuan bin Baharum is an Associate Professor who has served as a researcher and educator at Universiti Teknologi Malaysia for the past ten years. His area of specialization is Technology and Language Education, and his research focuses on language learning, psychology in education, and instructional technology, with recent work examining how self-transcendence, emotional intelligence, and enjoyment influence English learners.

Ali Hamed Barghi holds a PhD in Teaching English to Speakers of Other Languages (TESOL) and is currently a doctoral fellow in Curriculum and Instruction, specializing in teaching and teacher education, at Texas A&M University, USA. A certified TESOL teacher educator with extensive experience in teaching and teacher training, his research focuses on second language teacher education, teacher beliefs and agency, and transformative teacher education.

Seriaznita Mat Said is an Associate Professor who has served as an educator and researcher at Universiti Teknologi Malaysia for the past twenty years. Her areas of expertise include Teaching of English as a Second Language, English language learning and communication, English for Specific Purposes, and Discourse Analysis. To date, she has authored a research book, as well as indexed and non-indexed journal articles and book chapters and has served as a reviewer for several reputable indexed journals. She is currently the Head of the Research Group in Applied Linguistics and Language Teaching at Universiti Teknologi Malaysia. Since 2014, she has also served as an academic advisor to the English Program Unit at the Matriculation Centre of Islamic University Selangor. Her community service includes delivering motivational talks and designing short-term projects related to second language acquisition and assessment.

The Effect of Learning Modalities and Self-Regulated Learning on EFL Learners' Critical Thinking Development

Rosmania Rima*

Doctoral Program in Education, Universitas Sultan Ageng Tirtayasa, Serang, Indonesia

Yudi Juniardi

English Education Department, Faculty of Education, Universitas Sultan Ageng Tirtayasa, Serang, Indonesia

Syafrizal Syafrizal

English Education Department, Faculty of Education, Universitas Sultan Ageng Tirtayasa, Serang, Indonesia

Abstract—This research investigates the effect of learning modalities (online, blended, face-to-face) on the critical thinking skills of EFL learners, reviewed by their level of self-regulated learning (SRL), and develops conjectures about their interaction. Using a mixed-methods approach with an explanatory sequential design (quantitative: 3x3 factorial; qualitative: case study), it involved 99 fourth-semester learners enrolled in a Critical Reading course. The results revealed significant effects of both SRL level ($F = 16.663, p = 0.000$) and learning modalities ($F = 19.395, p = 0.000$) on critical thinking skills, as well as a significant interaction effect between the two ($F = 3.713, p = 0.008$), indicating that the effect of a learning modality depends on the level of SRL and vice versa. The subsequent qualitative analysis developed nine conjectures to explain this interaction, highlighting how high SRL learners benefit from adaptive independence in online learning, moderate SRL learners require scaffolded critical thinking, and low SRL learners face cognitive fragmentation barriers; how blended learning supports high SRL learners through trans-modality synergy, moderate SRL learners through integrated stimulation, and low SRL learners through adaptive instructional alignment; and how face-to-face learning fosters self-driven criticality in high SRL learners, provides focused stimulation for moderate SRL learners, and is constrained by cognitive discontinuity for low SRL learners. These findings carry significant pedagogical implications for language teaching, highlighting the need for adaptive learning strategies tailored to SRL levels to effectively promote both student autonomy and critical thinking across diverse learning environments.

Index Terms—blended learning, critical thinking, face-to-face learning, online learning, self-regulated learning

I. INTRODUCTION

The landscape of higher education is undergoing a profound transformation driven by rapid technological advancements, shifting pedagogy from passive knowledge transmission to student-centred engagement. This aligns with global demands and national policies, such as the Indonesian Minister of Education, Culture, Research, and Technology Regulation Number 53 of 2023, which mandates flexibility to foster lifelong learning. Such provisions necessitate integrating diverse modalities—face-to-face, online, and blended—to cultivate essential 21st-century skills like critical thinking and adaptability, which are crucial for navigating the digital era (van Laar et al., 2020).

The availability of these modalities presents distinct opportunities and challenges for EFL instruction. Online learning, accelerated by the pandemic, increases accessibility and flexibility, a trend evident at Sultan Ageng Tirtayasa University (Untirta) (Suharsih & Wijayanti, 2021). While digital resources can deepen understanding and produce outcomes comparable to those of traditional methods, the lack of immediate interaction may impede the verbal communication skills necessary for nuanced interpretation of EFL texts (Sari, 2021; Tahir, 2021). Conversely, Blended Learning has emerged to bridge this gap by integrating online autonomy with the depth of face-to-face interaction (Huang et al., 2022). Untirta supports this through its SPADA platform (Agung & Yuliana, 2024). Meanwhile, traditional face-to-face instruction remains vital for immediate feedback, collaborative analysis (Suhirman & Prayogi, 2023), and social engagement (Zhao & Song, 2022).

Despite the availability of these methods, the cultivation of critical thinking—defined as the ability to evaluate arguments and make evidence-based decisions (Ennis, 2011)—remains a challenge. Recent research at Untirta indicates that critical thinking skills among English Language Education students are primarily low to moderate (Rima & Muhyidin, 2023), underscoring the need for pedagogical intervention. The Critical Reading course serves as a pivotal platform for this development, training students to move beyond information acquisition to deep engagement with diverse perspectives

* Corresponding Author. Email: rosmania@untirta.ac.id

and complex arguments.

Crucially, the effectiveness of these modalities is likely influenced by Self-Regulated Learning (SRL). SRL refers to an individual's ability to control and adapt their learning processes (Schunk, 2023). It is theorised to moderate the relationship between instructional delivery and skill acquisition; for instance, highly self-regulated EFL learners can better leverage online independence (Lazorak et al., 2021). Research consistently highlights SRL's vital role in enhancing learning effectiveness and fostering critical thinking (Ranjbar et al., 2022; Zeng et al., 2024), yet its specific role in EFL Critical Reading remains underexplored.

While research on individual modalities exists, a comprehensive investigation into the differential effects of online, blended, and face-to-face learning on critical thinking in Critical Reading courses is limited (Ifnaldi et al., 2025). To address these gaps, this study employs a sequential explanatory mixed-methods design to evaluate the effectiveness of these three modalities and SRL levels on EFL learners' critical thinking. Specifically, this research aims to:

1. determine the differential main effects of learning modalities on critical thinking;
2. determine the differential main effects of SRL levels on critical thinking;
3. analyse the interaction effect between instructional delivery and learners' self-regulation capabilities;
4. qualitatively explore the underlying conjectures and specific self-regulation strategies that explain how learners navigate these learning environments to construct meaning.

II. LITERATURE REVIEW

A. Online Learning: Characteristics, Benefits, and Challenges

Online learning is a distance education method that leverages the Internet, enabling teaching and learning without spatial or temporal constraints (Jensen et al., 2022). Online learning encompasses various forms, including virtual interactions via video conferencing and Learning Management Systems (LMS) like Google Classroom, as well as fully online methods without direct physical interaction. Its primary benefits include flexibility in time and place, efficient dissemination of content, and unrestricted access to up-to-date information. Empirical studies indicate that these features can foster a more dynamic interaction between lecturers and students by facilitating real-time feedback and inclusive participation (Yu et al., 2022; Zeinstra et al., 2022). However, online learning also presents significant drawbacks, notably internet connectivity issues, reduced direct communication, a potential decrease in assessment objectivity, and adaptation challenges for students. To mitigate these challenges and standardise the quality of digital instruction, Untirta implemented the SPADA Online Learning System. SPADA is a comprehensive ecosystem designed to foster flexible, interactive learning. By utilising its interactive discussion forums and structured content modules, the platform supports the development of critical thinking and problem-solving skills. At Untirta, the SPADA Online Learning System has been implemented as an advanced platform to foster flexible and interactive learning, thereby promoting the development of critical thinking and problem-solving skills.

B. Blended Learning: Integration, Components, and Implementation

Blended learning is an educational approach that integrates traditional face-to-face instruction with online learning via digital platforms, offering students the flexibility to learn both in the classroom and independently online (Johler, 2022). By combining the strengths of both modalities, it serves as a natural extension of classroom learning, adapting to individual needs through interactive multimedia. Essential components for its implementation include live events, self-paced learning, collaboration, and assessment, with a focus on integrating direct interaction with independent online study (Al-Othman, 2025).

In the specific context of Critical Reading, this approach offers distinct advantages. Empirical studies indicate that the flexibility of blended learning allows students to process complex texts at their own pace before engaging in synchronous analysis (Boelens et al., 2017). This combination of online preparation and face-to-face debate fosters deeper reflection and increased student participation (Sivakumar et al., 2023). However, challenges such as gaps in digital literacy and inconsistent infrastructure persist (Listyani, 2025). To address these variables, Untirta mandates the use of the SPADA platform for the 2024/2025 course. This design integrates asynchronous modules for independent text analysis with synchronous video conferences and practical sessions, ensuring a comprehensive and supported learning experience.

C. Face-to-Face Learning: Foundations, Advantages, and Limitations

Face-to-Face Learning (FTFL) is a long-established and effective instructional method involving direct interaction between lecturers and students in a structured classroom environment (Gherheş et al., 2021). This method enables lecturers to observe student behaviour, provide personalised and timely feedback, and foster intensive social interaction. FTFL encompasses various types, including lectures, discussions, practical exercises, and group work, with lectures and interactive discussions being the most common. FTFL strategies can be either lecturer-centred (traditional) or student-centred (modern), with the latter aiming to develop students' affective and psychomotor domains (Ni, 2022). The benefits of FTFL include enhanced learning motivation, better comprehension of the material, effective two-way communication, flexibility without internet dependency, direct character assessment, and support for practical activities (Koirewoa et al., 2024). However, FTFL also has drawbacks, including fixed schedules that limit flexibility, challenges with learning

communities and physical presence, reliance on physical resources, uneven learning quality, student dependence on lecturers, time constraints, lack of lecture recordings, and classroom distractions (Juneja et al., 2023).

D. Critical Thinking Skills: Definition, Attributes, and Measurement

Critical thinking is recognised as a fundamental skill in higher education, involving the comprehensive analysis, evaluation, and synthesis of information (Ennis, 2015). To operationalise this concept, this study adopts Robert Ennis's Taxonomy of Critical Thinking. This framework distinguishes five specific categories of cognitive skills: (1) Elementary Clarification (focusing on a question), (2) Basic Support (judging credibility of sources), (3) Inference (making deductions and inductions), (4) Advanced Clarification (defining terms and assumptions), and (5) Strategies and Tactics (interacting with others). In this research, the five-stage framework serves as the primary measurement tool. The specific indicators within these categories were used to construct the scoring rubric for the Critical Reading tests, enabling a precise evaluation of how the blended learning intervention affected each dimension of students' analytical abilities.

E. Self-Regulated Learning: Conceptual Framework and Influencing Factors

Self-regulated learning (SRL) refers to an individual's capacity to actively control and manage their learning process, a concept rooted in Bandura's social cognitive theory, in which behaviour, personal factors, and the environment reciprocally influence learning (Zimmerman, 2023). This proactive approach involves setting clear goals, planning effective strategies, monitoring progress, and self-evaluating outcomes, emphasising the active role of metacognition, motivation, and behaviour. Factors influencing SRL include individual aspects such as knowledge (declarative, procedural, and conditional), metacognition (planning, organising, and evaluating), and goal-setting; behavioural factors, such as self-observation, self-judgment, and self-reaction, also play a crucial role, alongside environmental factors that shape learning behaviour and create conducive learning atmospheres (Henderson & Cunningham, 2023). This study operationalises SRL using Zimmerman's (2002) fourteen strategies as primary indicators: self-evaluation, organising, goal-setting/planning, seeking information, monitoring, environmental structuring, self-consequences, rehearsing, seeking assistance, and reviewing records. This framework is used to analyse whether the autonomy in blended learning encourages EFL learners to mobilise these strategies more effectively than in traditional settings.

III. METHODOLOGY

A. Design

This research employed a mixed-methods sequential explanatory design, characterised by the collection of quantitative data followed by qualitative analysis to elaborate on EFL students' performance (Creswell & Clark, 2017). This sequence leveraged complementary strengths: the quasi-experimental phase provided a robust empirical foundation, while the qualitative phase offered interpretive depth to contextualise statistical findings, including the exploration of outliers or unexpected patterns (Borg & Gall, 1989).

The quantitative phase utilised a 3x3 Factorial Design (Table 1). This specific configuration involves the crossing of two independent variables, each containing three distinct levels, yielding nine possible experimental conditions (Andrade, 2024). The design was selected to address practical constraints that prevented random assignment of individual students; instead, pre-existing intact classes were assigned to treatment groups. This structure enabled the simultaneous examination of the main effects and interaction effects of two independent variables: (1) Learning Modalities (three levels: online, blended, face-to-face) and (2) SRL Level (three levels: high, moderate, low) on the dependent variable, critical thinking skills.

TABLE 1
FACTORIAL DESIGN 3x3

Factor		Learning Modalities		
		Online (A1)	Blended (A2)	Face to Face (A3)
Self-regulated learning	High (B1)	B1/ A1	B1/ A2	B1/ A3
	Moderate (B2)	B2/ A1	B2/ A2	B2/ A3
	Low (B3)	B3/ A1	B3/ A2	B3/ A3

The qualitative phase employed a case-study approach to develop a substantive theory grounded in an in-depth understanding of the observed phenomena. Data were collected using in-depth semi-structured interviews to capture the learners' subjective experiences and reasoning processes. Following the data collection, the analysis involved three main stages: Open Coding, Axial Coding (application), and Selective Coding, designed to extract and formulate theoretical constructs from the empirical data (Vollstedt & Rezat, 2019).

B. Participants

Participants comprised 99 fourth-semester undergraduate students in the English Language Education Program at Untirta (2024/2025). Distributed across three intact classes (30–35 students each) undertaking the Critical Reading course, groups were assigned to distinct modalities: Online, Blended, and Face-to-Face, facilitating a 3x3 factorial design. Core objectives and materials remained consistent to ensure comparability, while delivery was adapted to the affordances of

the modality (asynchronous vs. synchronous discussions). Fourth-semester students were selected for their cognitive maturity and the course's curricular relevance. Subsequently, participants for the qualitative phase were purposively sampled based on critical thinking scores and SRL levels to ensure diverse representation.

C. Data Collection Instrument

Critical thinking skills were quantitatively assessed using a Critical Thinking Test administered as a pretest and posttest. Adapted from Ennis's (1985) framework, the instrument consisted of 12 essay-type items designed to evaluate the five specific cognitive dimensions previously outlined. To ensure content validity, the test items were reviewed by three senior lecturers in applied linguistics, and reliability was established with a Cronbach's alpha coefficient of 0.917.

Simultaneously, learners' SRL levels were measured using a comprehensive SRL Questionnaire comprising 55 items. This instrument was developed based on Zimmerman's (1986) self-regulated learning strategies, utilizing a 4-point Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree'. The questionnaire demonstrated strong internal consistency ($\alpha = 0.964$) and was used to categorize learners into high, moderate, and low SRL groups for the factorial design.

To gain a deeper understanding of these quantitative patterns, the study employed a purposive sampling strategy, specifically extreme-case sampling. From the total population, 27 participants were selected to represent the distinct intersections of the 3x3 design. This selection was limited to a small, focused cohort to enable the granular analysis required for a case study approach. Data were collected through semi-structured interviews, with questions adapted from Zimmerman's (1986) SRL interview protocols. The interview guide was designed to probe the specific cognitive conjectures and regulation strategies identified in the quantitative phase.

D. Data Analysis

(a). Quantitative Data Analysis

This research employed a mixed-methods approach, utilizing quantitative analysis to rigorously assess the research problem. Before conducting parametric tests, the fundamental assumptions were verified in SPSS v26. Results from the Kolmogorov-Smirnov and Levene's tests were non-significant ($p \geq 0.05$), confirming that the data met the necessary conditions of normality and homogeneity. A 3x3 between-subject factorial design was subsequently employed to examine the combined effects of learning modalities (online, blended, face-to-face) and SRL levels (high, moderate, low). A Two-way ANOVA was used to analyze main effects and interactions, with Scheffé's post hoc test used to identify specific group differences where significant interactions occurred at the 95% confidence level.

(b). Qualitative Data Analysis

Qualitative data, derived from questionnaires and interviews, were analysed using a case study. This approach aimed to analyse substantive theory through empirical data using focused case studies (Creswell, 2015). The analysis involved three main stages: Open Coding, Axial Coding, and Selective Coding. The initial stage, Open Coding, involved assigning codes to every participant's response from tasks and interviews. These codes were used to identify emerging ideas or concepts related to critical thinking and problem-solving. Following this, Axial Coding involved selecting a core category from the open coding results and linking other categories to it, specifically to describe the relationship between SRL and critical thinking. Key elements considered in axial coding included Causal Condition, Core Category, Context, Intervening Conditions, Strategies, and Consequences (Kendall, 1999). The final stage, Selective Coding, involved a comprehensive theory-building process that integrated and connected the various categories identified during axial coding.

IV. RESULTS AND DISCUSSION

A. Quantitative Findings

Methodological validity was ensured through parallel pre- and post-tests for Critical Thinking and a baseline SRL Questionnaire to stratify participants. Moderate SRL was the predominant category, and assumptions for parametric analysis were confirmed via Kolmogorov-Smirnov and Levene's tests ($p > 0.05$). The Two-way ANOVA (Table 2) revealed significant main effects on N-Gain for both Learning Modalities [$F(2, 90) = 19.395, p = 0.000$] and SRL [$F(2, 90) = 16, p = 0.001$]. Crucially, a significant interaction effect [$F(4, 90) = 3.713, p = 0.008$] indicates that instructional efficacy is contingent on learners' self-regulation, with the model explaining 48.3% of the variance ($R^2 = 0.483$).

TABLE 2
TWO-WAY ANOVA MEAN SCORE OF CRITICAL THINKING N-GAIN SCORE

Tests of Between-Subjects Effects					
Dependent Variable: N Gain Critical Thinking					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.462 ^a	8	.058	10.504	.000
Intercept	3.636	1	3.636	660.994	.000
Learning Modalities	.213	2	.107	19.395	.000
Self-Regulated Learning	.183	2	.092	16.663	.000
Learning Modalities * Self-Regulated Learning	.082	4	.020	3.713	.008
Error	.495	90	.006		
Total	4.631	99			
Corrected Total	.957	98			

a. R Squared = .483 (Adjusted R Squared = .437)

(a). *The Effect of Self-Regulated Learning on Learners' Critical Thinking*

To further explore the significant main effect of SRL, post hoc Scheffé tests (Table 3) revealed distinct differences. The comparison between High and Low SRL yielded a statistically significant difference ($p = 0.000$), with High SRL learners achieving a superior N-Gain (Mean Difference = 0.1153). Similarly, High SRL learners significantly outperformed the Moderate group ($p = 0.000$; Mean Difference = 0.0764). However, the difference between Moderate and Low SRL was not statistically significant ($p = 0.090 > 0.05$). Consequently, while High SRL learners demonstrated substantially better gains in critical thinking than both other groups, the performance of Moderate and Low SRL learners remained comparable.

TABLE 3
POST HOC TEST OF SELF-REGULATED LEARNING LEVEL FACTOR

Multiple Comparisons						
Dependent Variable: N_Gain_Critical Thinking						
Scheffe						
(I) Self_Regulated_Learning	(J) Self_Regulated_Learning	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
SRL_High	SRL_Moderate	.0764*	.01859	.000	.0301	.1227
	SRL_Low	.1153*	.01958	.000	.0665	.1640
SRL_Moderate	SRL_High	-.0764*	.01859	.000	-.1227	-.0301
	SRL_Low	.0389	.01749	.090	-.0047	.0824
SRL_Low	SRL_High	-.1153*	.01958	.000	-.1640	-.0665
	SRL_Moderate	-.0389	.01749	.090	-.0824	.0047

Based on observed means.

The error term is Mean Square (Error) = .006.

*. The mean difference is significant at the .05 level.

This study highlights the pivotal role of SRL in fostering critical thinking, particularly within critical reading. The ANOVA results confirm significant differences across SRL groups, corroborating meta-analyses (Broadbent & Poon, 2015) and frameworks that demonstrate that learners who actively monitor their cognitive processes outperform peers. This aligns with Zimmerman's (2002) cyclical phases of forethought, performance, and self-reflection, which are inherent to critical reading and enable High-SRL EFL learners to apply strategies effectively (Russell et al., 2022; Schunk, 2023). While High vs. Moderate distinctions were not always significant, potentially due to a "ceiling effect" or domain knowledge factors (Radovic & Seidel, 2024), the need to tailor instruction to learner autonomy remains clear.

Regarding learning modalities, the analysis powerfully positions Blended Learning as the most effective mode, particularly when combined with high SRL. The superior N-Gain in this group aligns with the 'Community of Inquiry' framework (Garrison & Kanuka, 2024), suggesting a unique synergy where learners utilize online autonomy for preparation and face-to-face sessions for social knowledge construction. In contrast, fully online environments, potentially hindered by 'Transactional Distance', showed less responsiveness to internal SRL variations. Similarly, traditional face-to-face learning yielded lower N-Gain, likely due to a lack of structured independent processing time. Ultimately, these findings confirm that integrating blended instruction with strong self-regulation is key to optimal development of critical thinking.

(b). *The Effect of Learning Modalities on Learners' Critical Thinking*

Post-hoc Scheffe analyses were also conducted to explore the significant main effect of learning modality on N-Gain Critical Thinking. This analysis specifically identified which pairs of learning modalities and SRL exhibited substantial differences in their effect on critical thinking. The results indicate that blended learning significantly outperformed both online and face-to-face learning. Specifically, blended learning resulted in significantly higher N-Gain Critical Thinking scores compared to online learning (Mean Difference = 0.0918, $p < 0.001$). Similarly, blended learning showed substantially greater gains in critical thinking than face-to-face learning (Mean Difference = 0.0871, $p < 0.001$).

In contrast, no statistically significant difference in average N-Gain Critical Thinking scores was found between online and face-to-face learning (Mean Difference = -0.0046, Sig. = 0.969, $p > 0.05$). This indicates that, within the context of this study, the effectiveness of online and face-to-face learning in fostering gains in critical thinking was comparable.

TABLE 4
POST HOC TEST ON LEARNING MODALITIES FACTOR

Multiple Comparisons Dependent Variable: N_Gain_Critical Thinking Scheffé						
(I) Learning_Modalities	(J) Learning_Modalities	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Online	Blended	-.0918*	.01786	.000	-.1362	-.0473
	Face to face	-.0046	.01845	.969	-.0506	.0413
Blended	Online	.0918*	.01786	.000	.0473	.1362
	Face to face	.0871*	.01858	.000	.0409	.1334
Face to face	Online	.0046	.01845	.969	-.0413	.0506
	Blended	-.0871*	.01858	.000	-.1334	-.0409
Based on observed means. The error term is Mean Square (Error) = .006.						
*. The mean difference is significant at the .05 level.						

This research confirms a significant effect of learning modalities on critical thinking ($p = 0.000$). Post-hoc Scheffé analysis (Table 4) revealed that blended learning significantly outperformed both online (mean difference = 0.0918, $p < 0.001$) and face-to-face instruction (mean difference = 0.0871, $p < 0.001$), while no significant difference was found between online and face-to-face learning ($p = 0.969$). The superiority of blended learning aligns with the Community of Inquiry framework (Akyol et al., 2009; Garrison & Vaughan, 2007) and recent studies (Amenduni & Ligorio, 2022; Teng et al., 2024), attributing its success to the integration of flexibility, interaction, and cognitive presence. In contrast, the comparable but lower effectiveness of online and face-to-face modes stems from their respective limitations: online learning heavily demands high SRL, while lecture-heavy face-to-face settings often fail to foster the deep, active reflection required for critical thinking, particularly for students with low SRL (Lim et al., 2020).

Crucially, a significant interaction effect was observed between learning modality and SRL ($F = 3.713$, $p = 0.008$), indicating that effectiveness is contingent upon student self-regulation. In online environments, N-Gain consistently declines with lower SRL, reflecting a heavy reliance on self-management (Dai et al., 2023; Jin et al., 2023). Conversely, blended learning proves most effective for high SRL learners, who achieve the highest gains by strategically integrating both modalities (Lee, 2024). Meanwhile, face-to-face instruction resulted in moderate N-Gain across all levels, appearing significantly less sensitive to SRL variations than the other modalities.

(c). Interaction Effect of Self-Regulated Learning and Learning Modalities on Learners' Critical Thinking

The results of the two-way ANOVA revealed a statistically significant interaction effect between the learning modality and learners' SRL levels on their critical thinking abilities. The significance level ($p < 0.05$) for this interaction term indicates that the effect of a particular learning modality on critical thinking outcomes varies by the learner's level of self-regulation. Conversely, the effect of a student's SRL level on their critical thinking was contingent upon the learning modality they experienced. This finding suggests that the effectiveness of a given learning modality was not uniform across all levels of student self-regulation.

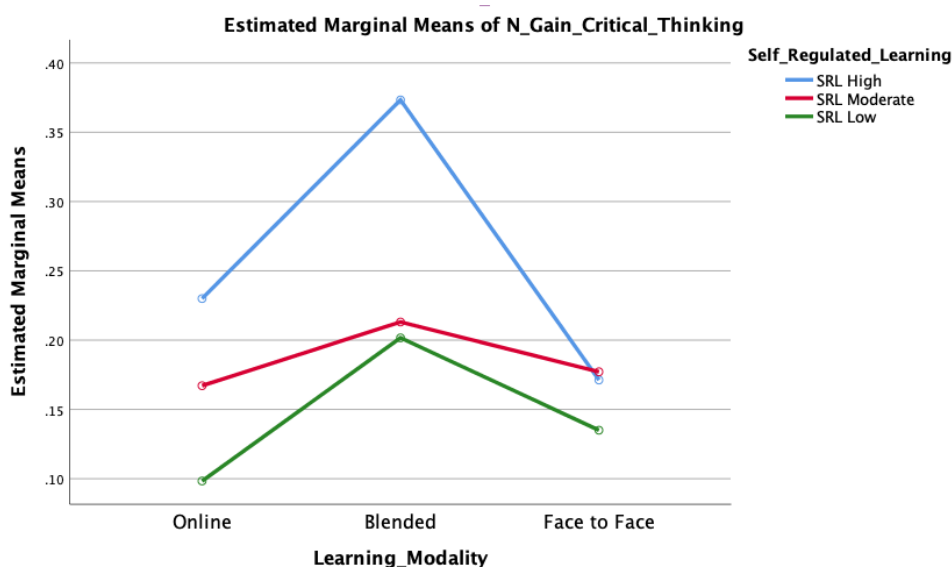


Figure 1. Interaction Effect of Self-Regulated Learning and Learning Modalities on Learners' Critical Thinking Skills

Figure 1 illustrates a disordinal interaction ($p = 0.008$), confirming that instructional effectiveness is contingent upon learners' SRL capacity. The overall model demonstrated a strong fit ($R^2 = 0.483$), explaining 48.3% of the variance in critical thinking scores. Scheffé post hoc analysis revealed that Blended Learning was superior, particularly for High SRL learners, who significantly outperformed their High SRL Online counterparts (Mean Diff = -0.1433 , $p = 0.044$). This advantage was even more pronounced against Moderate (Mean Diff = -0.2062 , $p < 0.001$) and Low SRL Online learners (Mean Diff = -0.2750 , $p < 0.001$). Conversely, within the online modality, SRL levels did not significantly differentiate gains, nor were significant differences found between Online and Face-to-Face groups across any levels ($p > 0.05$).

Blended Learning emerged as the most effective approach, particularly for High SRL EFL learners, who achieved a peak N-Gain of approximately 0.37. While scores decreased for Blended-Moderate (0.21) and Blended-Low (0.20) groups, they remained competitive. In contrast, Online Learning exhibited a sharp decline correlated with SRL capabilities, dropping from 0.17 (Moderate) to a nadir of 0.10 (Low), confirming heavy dependence on self-management. Face-to-Face instruction showed moderate stability, with a slight anomaly: Moderate SRL learners (0.18) marginally outperformed High SRL students (0.17), followed by Low SRL (0.14). Ultimately, while Face-to-Face and Online modes showed similar improvement levels, both fell short of the superior performance achieved by the Blended-High combination.

These interaction patterns provide strong empirical support for Adaptive Learning models (Dagunduro et al., 2024; du Plooy et al., 2024) and underscore the importance of instructional design that actively facilitates self-regulation (Vrieling et al., 2018). The efficacy of blended learning across levels suggests it successfully balances support and autonomy (Cao et al., 2024; Müller et al., 2023). Conversely, the evident challenges in fully online settings and the suboptimal performance in traditional contexts highlight the need for differentiated interventions (Jin et al., 2023; Prasse et al., 2024). Educators should therefore prioritize blended models and implement initial diagnostic SRL assessments to tailor strategies, ultimately optimizing gains in critical thinking in the Critical Reading course.

B. Qualitative Findings

To address the fourth objective regarding underlying success mechanisms, an in-depth qualitative analysis employed open and axial coding to formulate selective conjectures for each SRL-modality combination. The resulting axial coding diagram (Figure 2) visually maps the intricate framework connecting causal conditions, strategies, and consequences. These findings complement the quantitative analysis by providing nuanced explanations of how EFL learners with varying SRL levels engage with different learning environments to develop critical thinking.

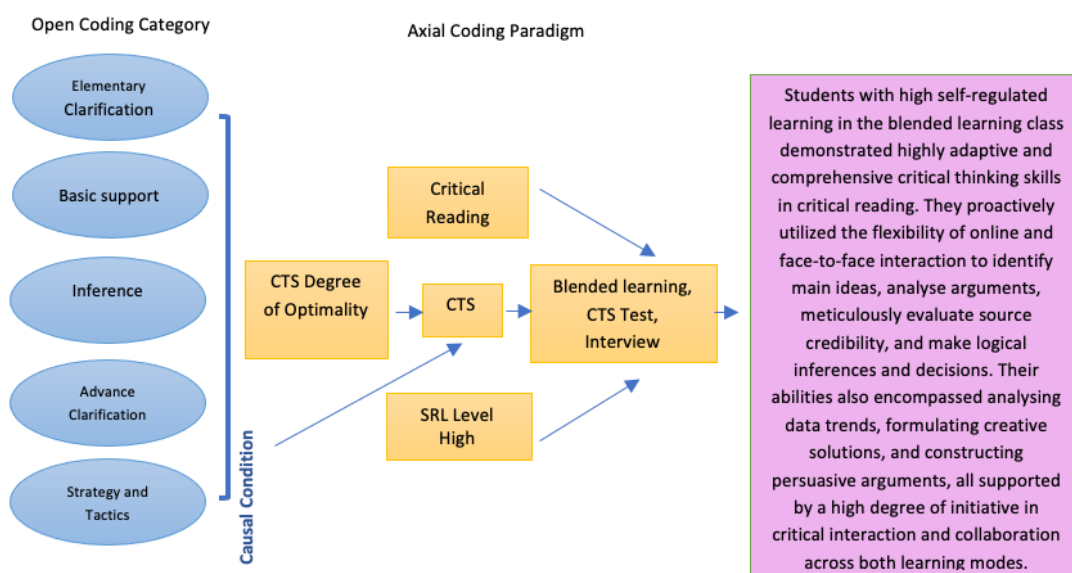


Figure 2. Axial Coding Diagram Illustrating the Relationship Between Learners With High Self-Regulated Learning (SRL) and Critical Thinking Skill in Blended-Learning

(a). High Self-Regulated Learning and Online Learning: Adaptive Independence

This conjecture asserts that high SRL EFL learners possess the essential foundational skills to thrive in the independent demands of online learning environments. Qualitative data indicate that their “adaptive independence”, where self-management and proactivity enable them to consciously and systematically utilize digital features and resources to sharpen their critical thinking abilities. These students efficiently focus on questions, accurately and comprehensively analyse arguments from varied online texts, and proactively seek clarification independently. They are also highly discerning in evaluating the credibility of online sources through cross-checking, capable of analysing digital data trends, and making logical and holistic deductive and inductive inferences. Even in the online context, they demonstrate creative

thinking in seeking alternative solutions, constructing persuasive arguments, and actively engaging in digital collaborations.

These findings align with the view that self-regulated learners possess metacognitive, motivational, and behavioural capabilities to achieve their learning goals (Brenner, 2022). Given the minimal external structure in EFL learning, this capacity becomes crucial. Highly SRL learners can apply effective planning, monitoring, and self-reflection strategies to navigate the vast amount of online information, identify key points, and critically evaluate the information they encounter. Self-regulated learning is a strong predictor of success in online learning, largely due to the self-paced, self-directed nature of the environment (Broadbent, 2017). The independence fostered by high self-regulated learning facilitates the complex process of online critical reading, where readers must actively interpret, analyse, and evaluate information without constant direct guidance.

(b). Moderate Self-Regulated Learning and Online Learning: Scaffolded Critical Thinking

This conjecture highlights that while moderate SRL EFL learners possess a sufficient base of self-regulatory skills (Wood et al., 2009), the development of their critical thinking abilities in online environments heavily relies on adaptive support and clear structure. The data suggest that their essential thinking abilities are reasonably developed, but their consistency and independent initiative fluctuate, necessitating external stimuli. These students can focus on questions about online English texts, primarily structured ones, but require more time for complex text analysis and explicitly analyze when they struggle to interpret their readings. They exhibit inductive-inductive or inductive-direct connections, but lack rigor, stating limitations. Creative thinking and persuasive argument construction are more frequent when explicitly prompted, and participation in online interactions is contingent on this.

These findings align with the scaffolding perspective in learning (Wood et al., 2009), which emphasizes the importance of temporary support to help learners navigate their zone of proximal development. Moderate SRL EFL learners are at a point where they possess basic capacity but require guidance and structure to consistently apply critical thinking strategies in a demanding online environment. Effective e-learning design often recommends features that support self-regulated learning, such as self-monitoring tools, clear instructions, and targeted feedback (Bozkurt & Sharma, 2021; Huang et al., 2022). In an online EFL critical reading class, this could involve providing guiding questions for critical reading, templates for argument analysis, or actively moderated discussion forums to stimulate deeper engagement and reflection. Without this adaptive support, learners with moderate SRL may not fully optimize their potential for critical thinking in online learning environments.

(c). Low Self-Regulated Learning and Online Learning: Cognitive Fragmentation

This conjecture demonstrates that the fundamental limitations in self-regulation among low-SRL EFL learners lead to “cognitive fragmentation” in an independent online environment. Qualitative data indicate that they struggle to focus on questions, often reading English texts superficially and being easily distracted. The understanding of analysis readability is minimal, and the analysis of online data is very superficial. Decisions tend to be impulsive, deductive-inductive reasoning is weak, and generalizations are often unwarranted. Difficulty defining terms and distinguishing concepts is also significant. Creative thinking is minimal, arguments are weak and incoherent, and participation in online interactions is passive.

This phenomenon aligns with findings that, without the intrinsic ability to manage time, self-motivate, or apply effective learning strategies, an online environment that demands autonomy becomes a barrier to learning (Pedrotti & Nistor, 2019). Cognitive Load Theory can explain this phenomenon: when low SRL learners are faced with complex online critical reading tasks without adequate scaffolding, the intrinsic and extraneous cognitive load becomes too high, making it difficult for them to process information and develop critical thinking skills (Hartson & Pyla, 2012). Limitations in self-regulated learning fundamentally hinder their ability to engage critically with the material, underscoring a need for significantly greater, structured, and proactive external support to lay a foundation for critical thinking.

(d). High Self-Regulated Learning and Blended Learning: Trans-Modality Critical Thinking Synergy

This conjecture asserts that High-SRL EFL learners strategically leverage blended learning to create a “trans-modality synergy” that optimizes critical thinking by proactively integrating the distinct advantages of both modes. Qualitative data reveal that they efficiently synthesize online resources with face-to-face clarifications to enhance argument analysis and systematically verify the credibility of sources. Their engagement is characterized by insightful questioning in class, followed by independent exploration in online forums. Consequently, they demonstrate strong deductive-inductive reasoning, creative thinking, and the ability to construct persuasive arguments, actively maintaining critical collaboration across both environments.

These findings align with the Community of Inquiry (CoI) framework, which emphasizes the maximization of cognitive, social, and teaching presence (Garrison & Vaughan, 2007). High-SRL EFL learners are ideal candidates for this model, as they demonstrate initiative to independently foster cognitive presence in online environments while leveraging face-to-face sessions to enhance social presence and deep interaction. By effectively combining online flexibility with the direct support of in-person learning, these students capitalize on the strengths of both modalities to achieve superior outcomes in the development of critical thinking (Lim & Yeo, 2021).

(e). Moderate Self-Regulated Learning and Blended Learning: Integrated Critical Stimulation

This conjecture suggests that moderate SRL EFL learners benefit greatly from “integrated critical stimulation” in blended learning to optimize the development of their critical thinking. Data indicate that integrating online and face-to-face modes fosters critical engagement, although their independent initiative and depth of analysis often depend on external triggers. They can focus questions on texts, guided by discussion, and analyse explicit arguments. They actively ask questions in face-to-face sessions and utilize online forums, though they are less proactive in seeking answers independently. Class discussions help them understand the criteria for evaluating sources, analyse complex data, and enrich their decision-making perspectives. Deductive-inductive reasoning develops with group discussion. They can define terms and distinguish concepts with face-to-face interaction. Creative initiatives are sparked by brainstorming, and logical arguments are developed through face-to-face debates or presentations. Interaction participation is active when collaborative tasks are assigned.

These findings are consistent with the notion that blended methods can provide flexible, adaptive scaffolding for learners who require it (Wilson, 2016). For moderate SRL EFL learners, face-to-face activities provide opportunities for direct guidance, immediate feedback, and social interaction that can trigger their critical thinking. At the same time, the online component guides them to practice independence, access additional resources, and reflect on the material without the pressure of time. In a Blended Learning method, this combination can overcome the weaknesses of single-mode learning, allowing instructors to provide appropriate support at the right time. An effective blended learning design can help students with moderate self-regulated learning internalize self-regulation and critical thinking strategies through repeated exposure and systematic feedback, thereby promoting significant improvement.

(f). Low Self-Regulated Learning and Blended Learning: Adaptive Instructional Alignment

This conjecture asserts that for low SRL EFL learners, blended learning necessitates “adaptive instructional alignment”. Fundamental limitations within this group lead to a heavy reliance on external support despite the availability of face-to-face interaction. Qualitative data reveal they struggle to optimize either modality due to a lack of initiative, requiring explicit direction to focus or analyze arguments. Specific observed deficits include passive questioning, minimal understanding of source credibility, and superficial observation of data without instructor interpretation. Furthermore, decision-making tends to be impulsive, while deductive and inductive reasoning, as well as creative thinking, remain weak, leading to passive participation in which concepts are not distinguished even after repeated explanations.

Although quantitative analysis suggests blended methods outperform pure face-to-face learning for this group, qualitative findings confirm that low self-regulation remains a significant impediment. These learners require highly intensive, personalized scaffolding; mere physical presence does not equate to active engagement without continuous prompting. This aligns with Van Laer and Elen (2017), who note that novice learners require highly structured instruction rather than the self-discovery approach often assumed in blended models. Consequently, the “self-directed” component poses a specific challenge, necessitating proactive pedagogical interventions to address SRL limitations before critical thinking strategies can be effectively internalized.

(g). High Self-Regulated Learning and Face-to-Face Learning: The Self-Driven Criticality

This conjecture posits that the initiative and intellectual independence of High-SRL EFL learners enable them to optimize critical thinking in face-to-face settings. Qualitative data indicate they actively shape discussions and extend understanding beyond classroom structures. Key behaviours include dissecting argument structures, posing challenging inquiries, and meticulously verifying the credibility of sources in English texts. Their reasoning is accurate, allowing them to question limitations, demonstrate creative thinking, and construct persuasive arguments. Consequently, they actively drive face-to-face interactions, initiating debates and facilitating in-depth collaboration.

These findings align with Self-Determination Theory (Ryan & Deci, 2020), describing these learners as active, autonomous learners who maximize the interactive potential of face-to-face EFL environments. Rather than merely absorbing information, they actively test understanding and challenge ideas through dialogue. Since critical thinking is a social and collaborative process (Liyanage et al., 2021), High-SRL learners leverage these interactions to hone skills beyond the explicit curriculum. By bringing a critical disposition to the classroom, they transform the face-to-face environment into an arena for sharpening their intellect.

(h). Moderate Self-Regulated Learning and Face-to-Face Learning: Focused Critical Thinking Stimulation

This conjecture suggests that while Moderate SRL learners possess basic self-regulation, their face-to-face critical thinking relies heavily on external structure, masking a lack of independent initiative. Engagement tends to be reactive and analysis superficial without direct guidance, limiting learners to safe boundaries rather than deep exploration of English texts. These findings align with Vygotsky’s Zone of Proximal Development (Iqbal, 2015), indicating that while scaffolding is necessary, non-adaptive support can constrain growth. Furthermore, consistent with Bandura’s (1986) social cognitive theory, a lack of motivation to actively extend observed behaviours prevents these learners from achieving the level of optimization seen in High SRL peers, as traditional EFL methods often fail to foster the autonomy required for advanced critical thinking.

(i). Low Self-Regulated Learning and Face-to-Face Learning: Cognitive Discontinuity

This conjecture asserts that fundamental limitations in self-regulation led to “Cognitive Discontinuity” among Low SRL EFL learners in face-to-face settings. Qualitative analysis indicates struggles with managing learning processes, manifesting as superficial reading of English texts, distraction, and passive participation. Argument analysis remains minimal and often incoherent, requiring step-by-step guidance. Furthermore, these learners avoid complex challenges, accept information uncritically, and exhibit impulsive decision-making, with weak deductive and inductive reasoning. Consequently, collaborative contributions are limited, marked by weak argumentation and a lack of constructive solutions.

These findings align with Cognitive Load Theory (Sweller, 2024), suggesting that the cognitive demands of face-to-face EFL environments are overwhelming for this group. Both extraneous and intrinsic loads hinder effective processing (Ouwehand et al., 2025), as the lack of metacognitive ability prevents independent knowledge building. It is emphasized that self-regulation is key to learning interactions; thus, without highly structured, explicit scaffolding, standard face-to-face methods become barriers to learning. This underscores the critical need for personalized interventions to stabilize essential thinking skills.

V. CONCLUSION

A. Conclusion

This research demonstrates the significant interactive effects of Self-Regulated Learning (SRL) and learning modalities on EFL learners’ critical thinking in Critical Reading. Quantitatively, higher SRL learners consistently outperformed their peers, while blended learning emerged as the superior modality. Crucially, a statistically significant interaction effect confirmed that instructional effectiveness is contingent upon EFL learners’ self-regulation capacity. The qualitative phase deepened these insights through selective conjectures, revealing how specific SRL mechanisms facilitate or hinder engagement across different environments. These findings corroborate the quantitative results and underscore the need for tailored instructional design, concluding that fostering critical thinking in EFL contexts requires integrating pedagogical approaches with individual students’ characteristics.

B. Implications

This research significantly enriches the theoretical understanding of how SRL and learning modalities interact within EFL education. It confirms SRL as a fundamental catalyst for critical thinking in critical reading, offering empirical evidence that its effectiveness is contingent on EFL learners’ characteristics rather than on a single universal mode. This supports adaptive learning frameworks and calls for instructional designs that explicitly integrate student self-regulation. Furthermore, the study contributes to blended learning literature by demonstrating its unique adaptability to various SRL levels, making it optimal for development through synergistic online and face-to-face interactions. In practice, these findings urge EFL curriculum designers and institutions to prioritize SRL development by adopting adaptive blended models that offer flexibility for high-SRL learners while providing structured scaffolding for moderate- and low-SRL learners. Diagnostic assessments are identified as crucial for tailoring this support. Consequently, EFL Critical Reading instructors should differentiate their pedagogy based on these diagnostics, explicitly teaching both critical thinking and self-regulation strategies. Ultimately, learners are encouraged to proactively develop their SRL skills to optimize their growth in critical thinking within these flexible environments.

C. Limitations and Future Directions

Several limitations are acknowledged. First, the single-institution setting restricts generalizability to broader EFL populations. Second, the self-developed SRL instrument relies on self-report, which creates potential for bias. Third, internal variations within modalities—such as interaction quality or structure—were not fully captured. Fourth, qualitative validity requires further cross-checking by a second coder. Finally, standard critical thinking metrics (post-test/N-Gain) may not fully capture the multidimensional complexity of thinking skills in a foreign language.

Future studies should employ longitudinal designs to track EFL learners’ development over time and to test specific pedagogical interventions for SRL. Since the current model explains 48.3% of the variance, future research must investigate the remaining 51.7% by exploring variables such as intrinsic motivation, emotional intelligence, and prior knowledge in EFL contexts. Additionally, researchers should aim for greater demographic diversity among EFL learners and develop more sensitive, context-specific instruments.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Pusat Pembiayaan dan Asesmen Pendidikan Tinggi (PPAPT), Ministry of Higher Education, Science, and Technology of Indonesia (Kemendikristek), for the Beasiswa Penyelesaian Studi Program Doktor Tahun 2025. This funding provided the necessary support for the completion and publication of this research.

REFERENCES

- [1] Adebukola, O. D., Chidinma, F. C., Olanike, A. A., & Ayo, A. E. (2024). Adaptive Learning Models for Diverse Classrooms: Enhancing Educational Equity. *International Journal of Applied Research in Social Sciences*, 6(9), 2228–2240. <https://doi.org/10.51594/IJARSS.V6I9.1588>
- [2] Afzal, A., Hager, G., Markidis, S., & Wellein, G. (2023). Making applications faster by asynchronous execution: Slowing down processes or relaxing MPI collectives. *Future Generation Computer Systems*, 148, 472–487. <https://doi.org/10.1016/J.FUTURE.2023.06.017>
- [3] Agung, A. M. B., & Yuliana, N. (2024). Pemanfaatan Teknologi Komunikasi SPADA UNTIRTA Dalam Meningkatkan Aktivitas Belajar. *CENDEKIA Jurnal Ilmu Sosial Bahasa Dan Pendidikan*, 4(1), 27–38. <https://doi.org/10.55606/cendikia.v4i1.2199>
- [4] Akyol, Z., Garrison, D., & Yasar, O. M. (2009). Online and blended communities of inquiry: Exploring the developmental and perceptual differences. *International Review of Research in Open and Distance Learning*, 10(6), 65–83. <https://doi.org/10.19173/IRRODL.V10I6.765>
- [5] Amenduni, F., & Ligorio, M. B. (2022). Blended Learning and Teaching in Higher Education: An International Perspective. *Education Sciences*, 12(2). <https://doi.org/10.3390/EDUCSCI12020129>
- [6] Andrade, C. (2024). Understanding Factorial Designs, Main Effects, and Interaction Effects: Simply Explained with a Worked Example. *Indian Journal of Psychological Medicine*, 46(2), 175–177. https://doi.org/10.1177/02537176241237066/SUPPL_FILE/SJ-XLSX-1-SZJ-10.1177_02537176241237066.XLSX
- [7] Boelens, R., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review*, 22, 1–18. <https://doi.org/10.1016/J.EDUREV.2017.06.001>
- [8] Borg, W. R., & Gall, M. D. (1989). *Educational Research*. New York: Longman. Retrieved June 1, 2025, from <https://www.scirp.org/reference/referencespapers?referenceid=648977>
- [9] Bozkurt, A., & Sharma, R. C. (2021). In Pursuit of the Right Mix: Blended Learning for Augmenting, Enhancing, and Enriching Flexibility. *Asian Journal of Distance Education*, 16(2). Retrieved June 10, 2025, from <https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/609>
- [10] Brenner, C. A. (2022). Self-regulated learning, self-determination theory, and teacher candidates' development of competency-based teaching practices. In *Smart Learning Environments* (Vol. 9, Issue 1). <https://doi.org/10.1186/s40561-021-00184-5>
- [11] Broadbent, J. (2017). Comparing online and blended learners' self-regulated learning strategies and academic performance. *Internet and Higher Education*, 33, 24–32. <https://doi.org/10.1016/j.iheduc.2017.01.004>
- [12] Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- [13] Cao, Y., Jeyaraj, J. J., & Bakar Razali, A. (2024). Challenges in Promoting Learner Autonomy in Blended Learning: Perspectives from English as a Foreign Language Teachers in China. *International Journal of English Language Education*, 12(2), 122. <https://doi.org/10.5296/IJELE.V12I2.22272>
- [14] Creswell, J. W., & Clark, V. L. P. (2017). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA Sage Publications. Retrieved June 1, 2025, from <https://www.scirp.org/reference/referencespapers?referenceid=2875785>
- [15] Creswell, J. W. (2014). *Research Design Qualitative, Quantitative and Mixed Methods Approaches (4th ed.)*. Thousand Oaks, CA Sage. Retrieved June 1, 2025, from <https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1964849>
- [16] Dai, Y., Luan, L., & Lin, X. (2023). The Effects of Online Learning Readiness on Self-Regulated Learning for the First-Time Online Learning Students. *Asian Journal of Distance Education*, 18(2), 42. <http://www.asianjde.com/>
- [17] du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21), e39630. <https://doi.org/10.1016/J.HELİYON.2024.E39630>
- [18] Ennis, R. (2011). Critical Thinking: Reflection and Perspective Part I. *Inquiry Critical Thinking Across the Disciplines*, 26(1), 4–18. <https://doi.org/10.5840/INQUIRYCTNEWS20112613>
- [19] Ennis, R. H. (2015). Critical Thinking: A Streamlined Conception. *The Palgrave Handbook of Critical Thinking in Higher Education*, 31–47. https://doi.org/10.1057/9781137378057_2
- [20] Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7, 95–105. <https://doi.org/10.1016/j.iheduc.2004.02.001>
- [21] Garrison, D. R., & Vaughan, N. D. (2007). Community of Inquiry and Blended Learning. *Blended Learning in Higher Education*, 13–30. <https://doi.org/10.1002/9781118269558.CH2>
- [22] Gherheș, V., Stoian, C. E., Fărcașiu, M. A., & Stanici, M. (2021). E-learning vs. Face-to-face learning: Analyzing students' preferences and behaviors. *Sustainability (Switzerland)*, 13(8). <https://doi.org/10.3390/SU13084381>
- [23] Hartson, R., & Pyla, P. S. (2012). UX Design Guidelines. In *The UX Book* (pp. 689–801). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-385241-0.00022-1>
- [24] Henderson, R. W., & Cunningham, L. (2023). Creating Interactive Sociocultural Environments for Self-Regulated Learning. In *Self-Regulation of Learning and Performance Issues and Educational Applications* (pp. 255–282). Routledge. <https://doi.org/10.4324/9780203763353-11>
- [25] Huang, M., Kuang, F., & Ling, Y. (2022). EFL learners' engagement in different activities of blended learning environment. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), 1–15. <https://doi.org/10.1186/S40862-022-00136-7/TABLES/6>
- [26] Ifnaldi, Juanda, Afnita, Yunus, A. F., & Afandi, I. (2025). Integrating Digital and Hybrid Learning Models in Environmental Literature Education: Enhancing Speaking Skills Through Virtual Platforms in the Digital Era. *Journal of Language Teaching and Research*, 16(5), 1682–1693. <https://doi.org/10.17507/JLTR.1605.26>
- [27] Iqbal, H. M. (2015). Cognitive Development, Educational Theories of. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, 51–57. <https://doi.org/10.1016/B978-0-08-097086-8.23099-7>

- [28] Jensen, J., Smith, C. M., Bowers, R., Kaloi, M., Ogden, T. H., Parry, K. A., Payne, J. S., Fife, P., & Holt, E. (2022). Asynchronous Online Instruction Leads to Learning Gaps When Compared to a Flipped Classroom. *Journal of Science Education and Technology*, 31(6), 718–729. <https://doi.org/10.1007/S10956-022-09988-7>
- [29] Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1), 1–21. <https://doi.org/10.1186/S41239-023-00406-5/TABLES/5>
- [30] Johler, M. (2022). Collaboration and communication in blended learning environments. *Frontiers in Education*, 7, 980445. <https://doi.org/10.3389/FEDUC.2022.980445/BIBTEX>
- [31] Juneja, P. K., Meadati, P. K., & Palreddy, S. R. (2023). A multi-attribute value approach for selection of course modality - A preliminary model. *IOP Conference Series: Earth and Environmental Science*, 1176(1). <https://doi.org/10.1088/1755-1315/1176/1/012010>
- [32] Kendall, J. (1999). Axial Coding and the Grounded Theory Controversy. *Western Journal of Nursing Research*, 21(6), 743–757. <https://doi.org/10.1177/019394599902100603>
- [33] Koirewoa, D. C., Cahyono, D., & Erwin. (2024). Effectiveness of online learning and face-to-face learning on students. *Indonesian Journal of Education (INJOE)*, 4(1), 249–264. <https://injoe.org/index.php/INJOE/article/view/118>
- [34] Laius, A., & Orgusaar, G. (2025). The Critical Role of Science Teachers' Readiness in Harnessing Digital Technology Benefits. *Education Sciences*, 15(8), 1001. <https://doi.org/10.3390/EDUCSCI15081001>
- [35] Lazorak, O., Belkina, O. V., & Yaroslavova, E. (2021). Changes in Student Autonomy via E-Learning Courses. *International Journal of Emerging Technologies in Learning*, 16(17), 209–225. <https://doi.org/10.3991/IJET.V16I17.23863>
- [36] Lee, C. M. (2024). Online Learning Versus Face to Face Learning toward Students: Which can be an Effective Way of Learning Methodology to our Current Educational System? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4854806>
- [37] Lim, C. L., Jalil, H. A., Ma'rof, A. M., & Saad, W. Z. (2020). Differences in Self-Regulated Learning (SRL) and Online Learning Satisfaction Across Academic Disciplines: A Study of a Private University in Malaysia. *International Journal of Learning*, 6(2), 62–67. <https://doi.org/10.18178/IJLT.6.2.62-67>
- [38] Listyani. (2025). Problems and Alternative Solutions of Indonesian EFL Teacher Trainees: A One-Decade Portrait of Online and Offline Classes. *Journal of Language Teaching and Research*, 16(2), 435–445. <https://doi.org/10.17507/JLTR.1602.09>
- [39] Müller, C., Mildenerger, T., & Steingruber, D. (2023). Learning effectiveness of a flexible learning study programme in a blended learning design: why are some courses more effective than others? *International Journal of Educational Technology in Higher Education*, 20(1), 1–25. <https://doi.org/10.1186/S41239-022-00379-X/METRICS>
- [40] Ni, L. B. (2022). Face-To-Face Learning, Teaching and Learning at Home Need to Go Hand in Hand. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 7(5), e001510. <https://doi.org/10.47405/MJSSH.V7I5.1510>
- [41] Ouwehand, K., Lespiau, F., Tricot, A., & Paas, F. (2025). Cognitive Load Theory: Emerging Trends and Innovations. *Education Sciences*, 15(4), 458. <https://doi.org/10.3390/EDUCSCI15040458>
- [42] Pedrotti, M., & Nistor, N. (2019). How Students Fail to Self-regulate Their Online Learning Experience. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11722 LNCS, 377–385. https://doi.org/10.1007/978-3-030-29736-7_28
- [43] Prasse, D., Webb, M., Deschênes, M., Parent, S., Aeschlimann, F., Goda, Y., Yamada, M., & Raynault, A. (2024). Challenges in Promoting Self-Regulated Learning in Technology-Supported Learning Environments: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/S10758-024-09772-Z>
- [44] Radovic, S., & Seidel, N. (2024). Self-regulated learning support in technology-enhanced learning environments: A reliability analysis of the SRL-S rubric. *International Journal of Assessment Tools in Education*, 11(4), 675–698. <https://doi.org/10.21449/IJATE.1502786>
- [45] Ranjbar Mohammadi, R., Saeidi, M., & Ahangari, S. (2022). The Effect of Self-regulated Learning Instruction on Reading Comprehension and Critical Thinking: PLS-SEM Approach. *Journal of School Psychology*, 11(2), 51–61. <https://doi.org/10.22098/JSP.2022.1713>
- [46] Rima, R., & Muhyidin, A. (2023). An Investigation of Critical Thinking Skills of Pre-Service English Teachers in The Context of Phonological Acquisition. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, 10(2), 229. <https://doi.org/10.33394/jo-elt.v10i2.9432>
- [47] Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- [48] Sari, S. I. (2021). Efektivitas Pembelajaran Daring dan Luring di SMP Negeri 3 Pleret. *Alinea: Jurnal Bahasa, Sastra, Dan Pengajaran*, 10(2), 145. <https://doi.org/10.35194/alinea.v10i2.1079>
- [49] Schunk, D. H. (2023). Self-Regulation of Self-Efficacy and Attributions in Academic Settings. In *Self-Regulation of Learning and Performance Issues and Educational Applications* (pp. 75–100). <https://doi.org/10.4324/9780203763353-4>
- [50] Sivakumar, A., Jayasingh, S., & Shaik, S. (2023). Social Media Influence on Students' Knowledge Sharing and Learning: An Empirical Study. *Education Sciences*, 13(7), 745. <https://doi.org/10.3390/EDUCSCI13070745>
- [51] Suharsih, S., & Wijayanti, M. A. (2021). Online Learning for EFL Learners: Perceptions, Challenges, and Expectations. *Journal of English Language Studies*, 6(2), 244–257. <https://doi.org/10.30870/JELS.V6I2.12122>
- [52] Suhirman, & Prayogi, S. (2023). Problem-based learning utilizing assistive virtual simulation in mobile application to improve students' critical thinking skills. *Humanities and Social Sciences Letters*, 11(3), 351–364. <https://doi.org/10.18488/61.V11I3.3380>
- [53] Sweller, J. (2024). Cognitive load theory and individual differences. *Learning and Individual Differences*, 110, 102423. <https://doi.org/10.1016/J.LINDIF.2024.102423>
- [54] Teng, Y., Yin, Z., Wang, X., & Yang, H. (2024). Investigating relationships between community of inquiry perceptions and attitudes towards reading circles in Chinese blended EFL learning. *International Journal of Educational Technology in Higher Education*, 21(1), 1–25. <https://doi.org/10.1186/S41239-024-00440-X/TABLES/9>

- [55] Thahir, R. (2021). Pengaruh Pembelajaran Daring Berbasis Google Classroom Terhadap Hasil Belajar Mahasiswa Pendidikan Biologi. *Edukatif: Jurnal Ilmu Pendidikan*, 3(4), 1936–1944. <https://doi.org/10.31004/edukatif.v3i4.1123>
- [56] van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2020). Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review. *SAGE Open*, 10(1). https://doi.org/10.1177/2158244019900176/ASSET/IMAGES/LARGE/10.1177_2158244019900176-FIG1.JPEG
- [57] Van Laer, S., & Elen, J. (2017). In search of attributes that support self-regulation in blended learning environments. *Education and Information Technologies*, 22(4), 1395–1454. <https://doi.org/10.1007/s10639-016-9505-x>
- [58] Vollstedt, M., & Rezat, S. (2019). An Introduction to Grounded Theory with a Special Focus on Axial Coding and the Coding Paradigm. In G. Kaiser & N. Presmeg (Eds.), *Compendium for Early Career Researchers in Mathematics Education* (pp. 81–100). Springer. https://doi.org/10.1007/978-3-030-15636-7_4
- [59] Vrieling, E., Stijnen, S., & Bastiaens, T. (2018). Successful learning: balancing self-regulation with instructional planning. *Teaching in Higher Education*, 23(6), 685–700. <https://doi.org/10.1080/13562517.2017.1414784>
- [60] Wilson, K. (2016). Critical reading, critical thinking: Delicate scaffolding in English for Academic Purposes (EAP). *Thinking Skills and Creativity*, 22, 256–265. <https://doi.org/10.1016/j.tsc.2016.10.002>
- [61] Wood, R. E., Beckmann, J. F., & Birney, D. P. (2009). Simulations, learning and real world capabilities. *Education and Training*, 51(5), 491–510. <https://doi.org/10.1108/00400910910987273>
- [62] Yu, H., Shi, G., Li, J., & Yang, J. (2022). Analyzing the Differences of Interaction and Engagement in a Smart Classroom and a Traditional Classroom. *Sustainability*, 14(13), 8184. <https://doi.org/10.3390/SU14138184>
- [63] Zeinstra, L., Kupers, E., Loopers, J., & De Boer, A. (2022). Real-time teacher-student interactions: The dynamic interplay between need supportive teaching and student engagement over the course of one school year. *Teaching and Teacher Education*, 121(2), 103906. <https://doi.org/10.1016/j.tate.2022.103906>
- [64] Zeng, W., Goh, Y. X., Ponnampuram, G., Liaw, S. Y., Lim, C. C., D/O Paramasivam, J., Chia, S. Y., & de Bruin, A. (2024). Promotion of self-regulated learning through internalization of critical thinking, assessment and reflection to empower learning (iCARE): A quasi-experimental study. *Nurse Education Today*, 142, 106339. <https://doi.org/10.1016/J.NEDT.2024.106339>
- [65] Zhao, S., & Song, J. (2022). Unpacking the Emotional Experiences of Learners in a Blended Learning Context. *Frontiers in Psychology*, 13(May), 1–10. <https://doi.org/10.3389/fpsyg.2022.879696>
- [66] Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. https://doi.org/10.1207/S15430421TIP4102_2/ASSET/CMS/ASSET/08E15D12-FD7A-4604-B4BE-D9F2DB261C66/S15430421TIP4102_2.FP.PNG
- [67] Zimmerman, B. J. (2023). Dimensions of Academic Self-Regulation: A Conceptual Framework for Education. In *Self-Regulation of Learning and Performance Issues and Educational Applications* (pp. 3–22). <https://doi.org/10.4324/9780203763353-1>

Rosmania Rima received her Master's degree in English Language Education from Muhammadiyah Prof. Dr. HAMKA University, Indonesia, in 2010. She serves as an Assistant Professor in the Faculty of Teacher Training and Education at Sultan Ageng Tirtayasa University (2008–Present), where she earned her Bachelor's degree in 2007 and is currently pursuing her Doctorate in Education. She has authored books such as *Reading Between the Lines: A Critical Reader's Guide* and *Case-Based Learning: The Path to EFL Learners' Critical Thinking*. Her research topics mainly include: English Language Teaching, Language Acquisition, Technology-Enhanced Language Learning, and Critical Thinking. E-mail: rosmania@untirta.ac.id ORCID iD: 0000-0002-5346-7945

Yudi Juniardi is a Professor in the Faculty of Teacher Training and Education at Sultan Ageng Tirtayasa University (Untirta). He earned a Bachelor's degree in English Education from Universitas Lampung (1997), and both a Master's degree in English Language Education (2007) and a Doctoral degree in Applied Linguistics (2012) from Universitas Negeri Jakarta (UNJ). His work in English Language Education has resulted in numerous international publications, presentations, and books, including *Research on ELT: From Theory to Practice* and *Pendidikan Karakter*; and memberships, including TEFLIN and AsiaTEFL. His research interests include Applied Linguistics, ICT in ELT, Reading, and Critical Thinking. E-mail: yudi.juniardi@untirta.ac.id ORCID iD: 0000-0003-3346-5288

Syafrizal Syafrizal earned a Doctoral degree in Language Education from Universitas Negeri Jakarta (UNJ) in 2012, having also completed a Diploma program at SEAMEO RELC, Singapore (2007) and a Sandwich Program at Ohio State University, USA (2010). He is now a Professor at the English Department, Faculty of Teacher Training and Education, Sultan Ageng Tirtayasa University. His work in English Language Teaching includes serving as the Head of International Office (2015–2019) and as a National Facilitator for the USAID DBE2 Program (2012–2017). His research interests include Teaching English as a Foreign Language (TEFL). E-mail: syafrizal@untirta.ac.id ORCID iD: 0000-0001-8671-1329

Saudi University EFL Students' Understanding of Plagiarism: Perceived Knowledge, Recognition Patterns, and Conceptual Ambiguity in Academic Writing

Nawal I. Alhodithi*

English Department, King Khalid University, Abha, Saudi Arabia

Abstract—This study examines Saudi university EFL students' understanding of plagiarism as a component of academic writing competence, with a focus on the relationship between perceived knowledge and actual recognition of plagiarism in common writing scenarios. Using a questionnaire adapted from Chu et al. (2020), 154 Saudi undergraduates enrolled in Research Writing at King Khalid University during the academic year 2024 responded to Likert-scale items measuring self-perceived understanding and scenario-based items assessing plagiarism recognition. Results revealed a notable discrepancy between confidence and competence. While a majority of students reported knowing how to avoid plagiarism ($M = 3.75$, $SD = 1.01$), their ability to identify plagiarism varied substantially across practices. High recognition rates were observed for explicit violations such as copying without citation (88.3%), but recognition declined sharply for practices requiring interpretive judgment, including paraphrasing without citation (53.2%) and AI-generated text submission (53.9%). Self-plagiarism (16.9%) and use of teacher-provided ideas without citation (18.2%) were largely unrecognized. These findings align with cross-cultural EFL research demonstrating that plagiarism understanding involves developmental and instructional dimensions beyond mere policy awareness. The study argues for integrating explicit, scenario-based plagiarism instruction into academic writing pedagogy to support EFL students in developing nuanced understandings of source use, attribution, and authorship in higher education contexts.

Index Terms—plagiarism, academic writing, EFL students, academic integrity, higher education

I. INTRODUCTION

Plagiarism is commonly treated as an issue of academic misconduct, framed in terms of rule violation and ethical breach. Universities typically address plagiarism through policies, detection software, and disciplinary procedures, operating on the assumption that students clearly understand what constitutes plagiarism and how to avoid it (Macfarlane et al., 2014; Park, 2003). However, research increasingly suggests that this assumption does not always hold, particularly for students writing in a second or foreign language (Pecorari & Petrić, 2014; Razi & Şahan, 2023). For many EFL students, plagiarism is not simply a matter of dishonest intent but is closely connected to challenges in academic writing, source integration, and understanding disciplinary expectations.

Academic writing requires writers to engage actively with sources, transform information, and position their own voices in relation to existing knowledge (Bazerman, 2013; Hyland, 2009). When students struggle to recognize what constitutes acceptable source use, their ability to contribute meaningfully to academic discourse is compromised. In this sense, understanding plagiarism is not only about avoiding academic penalties but about developing the skills necessary for legitimate knowledge construction. This perspective has gained increasing attention in applied linguistics research, which views plagiarism as a developmental issue embedded within academic literacy practices rather than a purely moral failing (Lea & Street, 2006; Pecorari, 2008).

These challenges are particularly salient in EFL contexts, where students often encounter English academic writing conventions for the first time at university (Mahmud & Mohd Tahir, 2024). Learning how to paraphrase, cite, and synthesize sources in a foreign language places additional cognitive and linguistic demands on students, making the boundaries between appropriate and inappropriate source use less transparent (Liu & Wu, 2020; Muluk et al., 2021). Research across East and Southeast Asian EFL contexts has consistently documented a significant mismatch between students' self-reported understanding of plagiarism and their actual performance in paraphrasing and source integration tasks (Hu & Lei, 2015; Liao & Tseng, 2010). This pattern suggests that EFL students may overestimate their ability to avoid plagiarism while lacking the procedural knowledge to apply their understanding in practice.

In the Saudi context, recent studies have begun to explore plagiarism perceptions among university students, though research specifically focused on EFL academic writing remains limited. Alsaedi and Alhumsi (2024) investigated Saudi

* Corresponding Author. Email: Nalhdiithi@kku.edu.sa

undergraduates' perceptions of plagiarism during e-learning sessions and found that most participants were aware of prevalent forms of plagiarism, though academic level, GPA, and gender did not significantly influence perceptions. Similarly, Alhadlaq et al. (2020) examined plagiarism attitudes among Saudi medical students and reported that attending ethics courses was associated with more negative attitudes toward plagiarism. Madkhali (2017) examined Saudi students studying in the United States and found that lack of knowledge about Western academic culture contributed to plagiarism challenges. More recently, Sheerah (2024) investigated digital academic dishonesty perceptions among Saudi EFL students at King Khalid University, highlighting issues including plagiarism, improper citation, and limited critical thinking skills. These studies indicate growing scholarly attention to academic integrity in Saudi higher education, yet gaps remain in understanding how Saudi EFL undergraduates perceive and recognize plagiarism within academic writing contexts.

Rather than focusing on instances of detected plagiarism, recent scholarship has begun to explore students' perceptions and understanding of plagiarism through scenario-based instruments that reveal conceptual ambiguity and gaps in academic literacy (Chu et al., 2020; Waltzer & Dahl, 2023). Building on this work, the present study examines Saudi EFL university students' understanding of plagiarism as part of their broader development as academic writers. By investigating both students' perceived knowledge and their ability to identify plagiarism across common writing scenarios, the study aims to contribute to ongoing discussions about academic writing instruction and integrity in Saudi EFL higher education.

II. LITERATURE REVIEW

A. *Plagiarism as an Issue of Academic Literacy*

Early discussions of plagiarism in higher education largely framed the issue within an ethics-based perspective, emphasizing misconduct, cheating, and institutional regulation (Park, 2003). From this viewpoint, plagiarism is primarily understood as a violation of academic norms that warrant detection and sanction. Subsequent reviews of academic integrity have reinforced this policy-oriented approach, highlighting the role of universities in maintaining standards and protecting scholarly credibility (Macfarlane et al., 2014).

While such perspectives are important for institutional governance, they offer limited insight into how students learn to write academically, particularly in a second language. Several scholars have argued that treating plagiarism solely as misconduct risks overlooking the complex learning processes involved in academic writing, especially for novice writers. Lea and Street (2006) proposed an academic literacies model that situates writing practices within broader social and disciplinary contexts, suggesting that students' source-use behaviors reflect their developing understanding of academic conventions rather than simple compliance or non-compliance with rules. Maharajh (2020), through focus-group data, demonstrated that students' perceptions of plagiarism are often shaped by uncertainty, contextual interpretation, and prior educational experiences rather than deliberate intent to deceive. More recently, Razi and Şahan (2023) demonstrated that integrating academic integrity instruction into EFL writing courses effectively reduced plagiarism and improved writing quality among secondary school students. These findings support a shift toward understanding plagiarism as part of academic literacy development.

B. *Plagiarism and Academic Writing in EFL Contexts*

Within applied linguistics, plagiarism has increasingly been examined in relation to second-language academic writing. Research on EFL writing consistently shows that students face significant challenges when integrating sources, paraphrasing complex texts, and maintaining authorial voice (Pecorari & Petrić, 2014). For EFL students, limited linguistic resources and unfamiliar rhetorical conventions can make it difficult to transform source material appropriately, increasing the likelihood of textual borrowing (Pecorari, 2003). Keck (2006) compared L1 and L2 writers' paraphrasing strategies and found that second-language writers relied more heavily on near-copy and minimal revision strategies, suggesting that linguistic constraints directly influence source transformation practices.

Liu and Wu (2020), studying Chinese EFL undergraduates, found that students often struggled to distinguish between acceptable paraphrasing and plagiarism, particularly when working with dense academic sources. Their study revealed that while students claimed awareness of the importance of paraphrasing and reported using summary, paraphrase, and quotation strategies when citing sources, their actual citation practices often fell short of academic expectations. Similarly, Muluk et al. (2021) reported that EFL students' plagiarism practices were frequently triggered by difficulties in understanding academic texts, time pressure, and insufficient instruction on source use. Li et al. (2024) examined how English academic writing textbooks in China address plagiarism and found limited text-based instruction, with paraphrasing strategies receiving emphasis through the concept of "indirect quotation." These studies highlight how plagiarism in EFL settings is closely tied to instructional practices and writing development rather than solely to ethical dispositions.

Research within diverse Asian EFL contexts has revealed consistent patterns in students' plagiarism understanding. For instance, Silfiani (2018) studied Indonesian undergraduates and concluded that students could distinguish between more and less serious forms of plagiarism. The students recognized that using another author's entire document represents a more severe violation than borrowing phrases. Students also identified acts such as copying statements without quotation marks, combining or changing statements without attribution, and submitting purchased assignments as plagiarism.

However, they did not consistently recognize that copying with quotation marks but improper citation, or translating sentences while acknowledging authors, could still constitute problematic source use. In the same vein, Hu and Lei (2015) examined Chinese university students and reported that cultural practices emphasizing imitation and memorization influenced students' conceptions of legitimate textual borrowing, with some students believing that plagiarism requires high textual similarity or complete duplication.

The concept of patchwriting further illustrates this developmental perspective. Howard (1995) introduced patchwriting to describe the process by which novice writers borrow and adapt language from source texts as they attempt to enter unfamiliar academic discourse. Later work by Pecorari (2008) emphasized that such practices should be understood as part of learning to write rather than automatic evidence of misconduct. Pecorari and Shaw (2012) showed that even advanced students struggle with intertextuality, further supporting the interpretation of source-use difficulties as developmental challenges. This distinction is crucial for interpreting EFL students' plagiarism judgments and practices within realistic educational contexts.

C. Students' Perceptions and Understanding of Plagiarism

Because plagiarism involves interpretation as well as rule knowledge, several researchers have focused on students' perceptions and self-reported understanding. Chu et al. (2020) demonstrated that students often overestimate their knowledge of plagiarism when compared with their actual ability to identify plagiaristic practices in specific scenarios. Their findings highlight a gap between confidence and competence, suggesting that self-perception alone is an unreliable indicator of plagiarism understanding.

A critical finding across EFL studies is this significant mismatch between students' self-reported understanding and their actual performance. Liao and Tseng (2010), examining Taiwanese postgraduate and undergraduate EFL students, found that despite participants denying having committed plagiarism and claiming awareness of the importance of paraphrasing, both groups failed to produce acceptable texts in paraphrasing tasks. Notably, undergraduates plagiarized more strings of words than graduate students, suggesting that cognitive development and metacognitive maturation contribute to improved performance over time. This gap indicates a lack of recognition of their own plagiarism or a fundamental misunderstanding of what constitutes proper paraphrasing.

Similar patterns have been observed in other higher-education populations. Shen and Hu (2021), in a mixed-methods study of Chinese graduate students, found that participants expressed strong awareness of plagiarism policies but demonstrated inconsistent judgments when evaluating real academic scenarios. Tran et al. (2022) likewise reported that postgraduate students' understanding of plagiarism varied widely depending on educational background and prior instruction, reinforcing the importance of academic socialization. Scenario-based approaches have proven particularly effective in revealing these conceptual ambiguities. Waltzer and Dahl (2023) showed that students' evaluations of plagiarism varied significantly depending on text type, authorship context, and instructional cues, suggesting that understanding plagiarism is not static but highly context-dependent.

D. Factors Influencing Plagiarism Understanding

Research has identified multiple factors influencing EFL students' plagiarism understanding and behaviors. Educational factors include lack of explicit instruction in paraphrasing, unsuccessful transfer of paraphrasing knowledge due to limited practice, lack of training in academic writing and source attribution, and unclear expectations about plagiarism (Hu & Lei, 2015; Liao & Tseng, 2010; Silfiani, 2018). Practical factors encompass time pressure, language difficulties, and limited English proficiency, leading to reliance on copying. Cultural factors play a significant role, particularly the influence of citation practices that differ from Western academic norms and historical traditions that may encourage imitation without attribution (Farahat, 2022). Institutional factors include a lack of deterrence policies and punitive rather than educative approaches to plagiarism (Hu & Lei, 2015).

The increasing use of text-matching software has further shaped how students understand plagiarism. While tools such as Turnitin are intended to support academic integrity, research suggests that students often interpret similarity scores as definitive indicators of ethical writing rather than as aids for learning (Alua et al., 2023; Nketsiah et al., 2024). This instrumental view may obscure deeper issues of source integration and paraphrasing competence. More recently, the emergence of generative AI tools has added new complexity to plagiarism recognition, underscoring the urgency of helping students develop principled understandings of originality, attribution, and responsible source use (AbdAlgane et al., 2025; Cotton et al., 2023). In the Saudi context, Alsofyani and Barzanji (2025) found that while ChatGPT-generated feedback improved EFL learners' writing skills, concerns about plagiarism and over-reliance on AI persist among students and educators.

E. Plagiarism Research in the Saudi EFL Context

Research on plagiarism perceptions among Saudi students has grown in recent years, though studies specifically focusing on EFL academic writing contexts remain limited. Madkhali (2017) examined Saudi students studying in the United States and found that unfamiliarity with Western academic conventions contributed to plagiarism challenges. Alhadlaq et al. (2020) investigated attitudes toward plagiarism among Saudi medical students and found that attending ethics courses was associated with more negative attitudes toward plagiarism, while students with higher GPAs also demonstrated greater awareness. Issrani et al. (2021) similarly found that Saudi students displayed adequate knowledge

of plagiarism, though awareness increased with academic level.

More recently, Alsaedi and Alhums (2024) examined Saudi undergraduate students' perceptions of plagiarism during e-learning sessions and found that most participants were aware of prevalent forms of plagiarism in EFL research writing, though academic variables did not significantly influence perceptions. Sheerah (2024) investigated perceptions of digital academic dishonesty among Saudi EFL students at King Khalid University and found that students engaged in various forms of academic dishonesty, including plagiarism, copy-paste behaviors, and inadequate citation practices, often due to limited proficiency with digital tools and academic writing conventions. AbdAlgane et al. (2025) compared AI-generated and human-written academic texts among Saudi EFL learners at Qassim University, finding significant implications for academic integrity and writing instruction.

These studies indicate growing attention to academic integrity in Saudi higher education, yet important gaps remain. Most existing research has focused on medical or general student populations rather than specifically on EFL learners developing academic writing competence. Furthermore, few studies have employed scenario-based approaches to examine not only what students report knowing about plagiarism but also how they actually recognize plagiarism across different types of academic writing practices.

F. Identified Gap

Taken together, the literature points to a need for research that examines plagiarism understanding as part of academic writing development, particularly in Saudi EFL university contexts. Although previous studies have explored students' perceptions of plagiarism in Chinese, Indonesian, and Taiwanese settings, there remains limited empirical evidence extending scenario-based approaches, such as the instrument developed by Chu et al. (2020), to Saudi EFL higher-education writing contexts. Moreover, existing research has paid relatively little attention to how different academic writing practices vary in their conceptual transparency to learners.

While many studies document EFL students' general awareness of plagiarism and reliance on self-reported understanding, plagiarism recognition is often treated as a unified construct. This approach risks obscuring important distinctions between practices that involve clear rule violations (e.g., copying without citation) and those that require interpretive judgment and academic literacy skills (e.g., paraphrasing, reuse of prior work, or AI-assisted writing). As a result, specific areas of conceptual ambiguity that have direct implications for academic writing instruction remain underexamined.

The present study addresses this gap by examining patterns of plagiarism recognition among Saudi EFL undergraduates across different categories of academic writing practices. By examining EFL students' responses to a validated, scenario-based questionnaire adapted from Chu et al. (2020), this study provides empirical insight into how Saudi EFL undergraduates understand plagiarism as an aspect of academic writing competence. Specifically, the study explores students' perceived knowledge of plagiarism, their reported ability to avoid plagiarizing behaviors, and their capacity to identify plagiarism across common academic writing scenarios involving source use, paraphrasing, authorship, and the use of external tools. By identifying which practices are consistently recognized, contested, or largely unrecognized, the study offers a nuanced account of plagiarism understanding that can inform differentiated instruction in academic writing pedagogy and support for Saudi EFL learners in higher-education writing contexts.

III. METHODOLOGY

A. Research Questions

The study addresses the following research questions:

- RQ1. How do Saudi EFL undergraduates perceive their understanding of plagiarism and their ability to avoid plagiarizing behaviors in academic writing?
- RQ2. How do Saudi EFL undergraduates identify plagiarism across common academic writing scenarios involving source use, paraphrasing, authorship, and the use of external tools?
- RQ3. Which academic writing practices show the greatest level of conceptual ambiguity in Saudi EFL undergraduates' identification of plagiarism?

B. Research Design

This study adopted a quantitative, survey-based research design to examine EFL university students' understanding of plagiarism within academic writing contexts. Rather than focusing on reported instances of misconduct, the study emphasized students' recognition and conceptual understanding of plagiarism, treating it as an aspect of academic writing competence. A scenario-based questionnaire was used to capture both students' perceived knowledge of plagiarism and their ability to identify plagiarism across common academic writing situations.

C. Participants

The participants were 154 female EFL undergraduate students enrolled in a research writing course at a public university in Saudi Arabia. The course focused on developing students' academic writing skills, including source use, paraphrasing, citation, and the ethical integration of references. English was used as the primary language of instruction for academic writing tasks. The all-female sample reflects the institutional structure of the university, where male and

female students are taught in separate sections. Participants represented different academic levels within the undergraduate program and were enrolled across more than one academic term. Participants were selected through convenience sampling and represented different academic levels within the undergraduate program. Participation in the study was voluntary, and students were informed that their responses would be used for research purposes only and would not affect their course grades. Ethical approval was obtained from the university's research ethics committee, and all participants provided informed consent before completing the questionnaire.

D. Instrument

Data were collected using a questionnaire adapted from Chu et al. (2020), originally developed to examine students' self-perception and actual understanding of plagiarism among secondary school students in Hong Kong. The instrument was selected because of its focus on plagiarism recognition through realistic academic scenarios, which aligns with the present study's emphasis on academic writing competence rather than policy awareness alone. To ensure appropriateness for Saudi EFL undergraduates, minor wording adaptations were made to ensure clarity and contextual appropriateness. These adaptations primarily involved simplifying lexical items and replacing culturally specific references with academically equivalent examples familiar to the participants. No changes were made to the core constructs, structure, or intent of the original instrument.

The questionnaire consisted of two main sections. The first section, addressing perceived understanding of plagiarism, included two Likert-scale items designed to measure students' self-reported confidence in their ability to recognize plagiarism and avoid plagiarizing behaviors in academic writing. Students indicated their level of agreement with each statement using a five-point scale ranging from strongly disagree (1) to strongly agree (5). The second section presented students with seven academic writing scenarios reflecting common practices in university-level writing. These scenarios addressed issues such as paraphrasing source material, reusing previous work, authorship boundaries, and the use of external tools. Students were asked to judge whether each scenario constituted plagiarism using the same five-point Likert scale. This section was intended to assess students' conceptual understanding and recognition, rather than their knowledge of institutional rules.

The questionnaire was piloted with 20 students before the main administration to check for clarity and comprehension. Based on pilot feedback, minor adjustments were made to the wording. The scenario-based items retained their original focus on recognition of plagiarism across common academic writing practices, allowing for meaningful comparison with findings reported in previous studies using the same instrument.

E. Procedures

The questionnaire was administered during the academic term in which participants were enrolled in Research Writing. Data were collected either during class time or through an online survey platform, depending on course scheduling. All students received the same instructions and completed the questionnaire individually. Before participating, students were informed about the purpose of the study and assured that their responses would remain anonymous and confidential. Completing the questionnaire indicated informed consent. No identifying information was collected.

F. Data Analysis

Data analysis focused on providing a clear and descriptive account of students' understanding of plagiarism, aligned with the study's research questions. Given that the research questions were descriptive in nature (examining how students perceive and identify plagiarism rather than testing hypotheses about relationships between variables), descriptive statistics were deemed appropriate. Responses to the Likert-scale items were analyzed using frequencies, percentages, means, and standard deviations to summarize students' perceived knowledge and confidence regarding plagiarism. For the scenario-based items, response patterns were examined to identify areas of agreement and disagreement among participants. Following Chu et al. (2020), responses of "Agree" or "Strongly Agree" were coded as recognition of plagiarism, while "Disagree" or "Strongly Disagree" responses were coded as non-recognition. Particular attention was given to scenarios that produced inconsistent or divided responses (i.e., recognition rates near 50%), as these were interpreted as indicators of conceptual ambiguity in students' understanding of plagiarism practices. All analyses were conducted using SPSS version 28.

Because the scenario-based items represented independent academic writing practices rather than indicators of a single latent construct, internal consistency reliability analysis was not applied to this section of the instrument. Each scenario assessed a distinct type of plagiarism-related practice (e.g., copying without citation, paraphrasing without attribution, self-plagiarism), and students' recognition of one practice would not necessarily predict their recognition of another. Instead, responses were analyzed descriptively to identify patterns of agreement and conceptual ambiguity across scenarios, in line with the analytical approach adopted by Chu et al. (2020).

IV. RESULTS AND DISCUSSION

A. Students' Perceived Understanding of Plagiarism

Analysis of the Likert-scale items indicated moderate to high variation in students' self-reported understanding of plagiarism. For the item "I am good at identifying acts of plagiarism," students reported a moderate level of confidence

($M = 3.08$, $SD = 0.89$, $N = 154$). Nearly half of the respondents (47.6%) selected the neutral option, while fewer than one-third (27.9%) agreed that they were good at identifying plagiarism. Only 4.1% strongly disagreed, and no students selected “strongly agree.” Despite the moderate mean score, the relatively large standard deviation and high proportion of neutral responses suggest that students’ confidence in identifying plagiarism was tentative rather than assured.

In contrast, responses to the item “I know how to avoid plagiarizing behaviors” reflected a higher level of perceived confidence ($M = 3.75$, $SD = 1.01$, $N = 154$). More than two-thirds of the participants (71.4%) agreed or strongly agreed that they knew how to avoid plagiarism, with 50.6% agreeing and 20.8% strongly agreeing. Relatively few students expressed disagreement (11.0% disagree; 3.2% strongly disagree), while 14.3% remained neutral. These results suggest that students felt more confident in their ability to avoid plagiarism than in their ability to clearly identify it, a pattern consistent with findings from Chu et al. (2020).

The combined mean of perceived understanding items was 3.42 ($SD = 0.95$), indicating moderate overall confidence in plagiarism knowledge among participants. Table 1 presents the descriptive statistics for students’ perceived understanding of plagiarism.

TABLE 1
DESCRIPTIVE STATISTICS FOR PERCEIVED UNDERSTANDING OF PLAGIARISM

Item	n	Mean	SD
I am good at identifying acts of plagiarism	154	3.08	0.89
I know how to avoid plagiarizing behaviors	154	3.75	1.01
Combined mean		3.42	0.95

B. Identification of Plagiarism Across Academic Writing Scenarios

Students’ responses to the scenario-based items revealed notable variation in their ability to recognize plagiarism across different academic writing practices. High levels of agreement were observed for scenarios involving explicit forms of plagiarism. Copying others’ work without citation was recognized as plagiarism by 88.3% of participants, while submitting another person’s work as one’s own was identified by 76.0%. These findings suggest that participants were generally able to identify clear violations of academic integrity.

In contrast, scenarios representing less explicit writing practices produced substantially lower recognition rates. Only 53.2% of students identified paraphrasing without citation as plagiarism, while recognition of submitting AI-generated work as one’s own was similarly divided (53.9%). Even lower recognition rates were observed for using uncited internet ideas (29.2%), using teacher-provided ideas without citation (18.2%), and reusing one’s own previous work without citation (16.9%). These patterns indicate that while students demonstrated awareness of plagiarism in principle, translating that awareness into accurate recognition within real academic writing contexts presented substantial difficulty. Table 2 summarizes the recognition rates across all seven scenarios.

TABLE 2
RECOGNITION OF PLAGIARISM ACROSS ACADEMIC WRITING SCENARIOS ($N = 154$)

Scenario	Recognition (%)
Copying others' work without citation	88.3
Submitting another person's work as one's own	76.0
Submitting AI-generated work as one's own	53.9
Paraphrasing without citation	53.2
Using uncited internet ideas	29.2
Using teacher-provided ideas without citation	18.2
Self-plagiarism (reusing own work)	16.9
Total/Mean recognition rate	47.9

C. Areas of Conceptual Ambiguity

The greatest conceptual ambiguity was observed in scenarios involving paraphrasing and AI-assisted writing, both of which generated near-equal agreement and disagreement among participants. With recognition rates of 53.2% and 53.9% respectively, these practices produced responses closest to a 50-50 split, indicating substantial uncertainty in students’ judgments. This suggests difficulty in applying plagiarism concepts to complex writing situations that require interpretation rather than rule recall.

Practices such as self-plagiarism and the use of instructional materials without citation were largely not recognized as plagiarism, with recognition rates below 20%. Rather than representing ambiguity, these low rates indicate persistent gaps in understanding of authorship and attribution norms in academic writing contexts. Students appeared not to view the reuse of their own work or teacher-provided ideas as requiring attribution, reflecting limited awareness of academic conventions regarding originality and intellectual ownership.

D. Discussion

The findings of this study highlight a notable discrepancy between Saudi EFL undergraduates’ perceived understanding of plagiarism and their actual ability to recognize plagiarism in academic writing scenarios. While students reported relatively high confidence in their ability to avoid plagiarizing behaviors ($M = 3.75$), their scenario-based judgments

revealed substantial uncertainty, particularly for practices requiring interpretive judgment. This mismatch mirrors the pattern observed by Chu et al. (2020), who similarly found that students' self-perceptions do not consistently align with their applied understanding of plagiarism. The findings also align with cross-cultural EFL research demonstrating that students tend to overestimate their paraphrasing abilities while failing to recognize their own problematic textual borrowing practices (Liao & Tseng, 2010; Liu & Wu, 2020).

The contrast between confidence and recognition is particularly evident in the scenario results. Although a large majority of participants correctly identified explicit cases such as copying others' work without citation, recognition rates declined sharply when students were asked to evaluate practices that required interpretive judgment. Only approximately half of the participants identified paraphrasing without citation and submitting AI-generated text as one's own work as plagiarism. These findings suggest that students' understanding of plagiarism is strongest when violations are overt but weakens substantially in situations that involve transforming or mediating sources. This pattern echoes Silfiani's (2018) finding that Indonesian EFL students could recognize serious plagiarism involving whole documents but showed uncertainty about less explicit forms of textual borrowing.

The low recognition rates observed for self-plagiarism (16.9%) and the use of teacher-provided ideas without citation (18.2%) further underscore conceptual gaps in students' understanding. Many students appeared not to view the reuse of their own work or instructional materials as requiring attribution, reflecting limited awareness of authorship and ownership conventions in academic discourse. Similar patterns have been reported in EFL contexts across Asia, where students' prior educational experiences may not emphasize citation as a rhetorical practice linked to knowledge construction (Hu & Lei, 2015; Muluk et al., 2021; Pecorari, 2008). These findings align with recent Saudi research; Alsaedi and Alhumsi (2024) similarly found that while Saudi undergraduates demonstrated awareness of explicit plagiarism forms, recognition of more nuanced practices varied. The findings also correspond with Sheerah's (2024) observation that Saudi EFL students at King Khalid University engaged in various forms of academic dishonesty partly due to limited understanding of proper citation practices.

The discrepancy between perceived and actual abilities can be attributed to multiple interacting factors identified in the literature. Educational factors, including lack of explicit instruction in paraphrasing and unsuccessful transfer of knowledge due to limited practice, appear particularly relevant to the present findings. As Liao and Tseng (2010) noted, students may understand plagiarism conceptually but lack the procedural skills to apply this knowledge, a pattern reflected in the present participants' high confidence in avoidance coupled with inconsistent recognition. Mahmud and Mohd Tahir (2024) demonstrated that educational interventions focusing explicitly on referencing skills can effectively improve ESL students' academic writing while reducing plagiarism. The institutional environment may also play a role; when institutions adopt punitive rather than educative approaches to plagiarism, students may lack opportunities to receive clear feedback about their actual performance and develop more accurate self-assessments.

Taken together, these findings support a view of plagiarism as an academic literacy issue rather than solely a question of ethical intent. The areas of uncertainty identified in this study correspond closely to linguistic and rhetorical challenges documented in research on EFL academic writing, particularly difficulties with paraphrasing, synthesis, and source integration (Keck, 2006; Liu & Wu, 2020). In this sense, students' misrecognition of plagiarism reflects gaps in writing instruction rather than deliberate wrongdoing. The developmental dimension is also apparent: just as Liao and Tseng (2010) found that undergraduates plagiarized more than graduate students, the present findings suggest that students at different stages of academic socialization may demonstrate varying levels of competence in recognizing and avoiding plagiarism.

The emerging ambiguity surrounding AI-generated text further highlights the evolving nature of academic authorship. The near-equal division in students' judgments suggests that institutional guidelines alone may not be sufficient to clarify expectations in technologically mediated writing contexts. AbdAlgane et al. (2025) found that while AI tools can improve linguistic features in Saudi EFL students' writing, concerns about over-reliance and diminished authentic writing ability persist. Without explicit instruction on how AI tools intersect with academic integrity, students may continue to struggle to apply plagiarism concepts consistently. This finding underscores the urgency of addressing new forms of textual production within academic writing pedagogy.

V. CONCLUSION

This study examined Saudi EFL undergraduates' understanding of plagiarism through a combination of self-reported perceptions and scenario-based judgments of academic writing practices. The findings demonstrate that although students expressed relatively high confidence in their ability to avoid plagiarism, their actual recognition of plagiarism varied systematically across different types of writing practices. Clear and explicit violations, such as copying others' work without citation, were consistently identified, whereas practices requiring interpretive judgment, particularly paraphrasing, self-plagiarism, the use of teacher-provided material, and AI-assisted writing, were characterized by substantial uncertainty or misrecognition.

By disaggregating plagiarism recognition across distinct categories of academic writing practices, the study moves beyond a binary conception of plagiarism understanding and highlights patterned zones of conceptual difficulty among Saudi EFL learners. These patterns suggest that plagiarism understanding develops unevenly and is closely tied to students' academic literacy development rather than merely to policy awareness or ethical intent. In this respect, the findings

support applied-linguistics perspectives that view plagiarism as a pedagogical issue embedded within academic writing instruction, especially for students writing in a second or foreign language.

The study also underscores the evolving nature of plagiarism in contemporary academic contexts. The divided judgments surrounding AI-generated writing indicate that emerging technologies have further blurred traditional notions of authorship and originality. Without explicit instructional guidance, students may struggle to integrate new tools into their academic practices in ways that align with institutional expectations. Addressing such challenges requires pedagogically informed approaches that situate plagiarism within broader discussions of responsible source use, attribution, and knowledge construction.

Overall, the findings reinforce the need for academic writing instruction that engages students with realistic writing scenarios and differentiates between types of plagiarism-related practices according to their cognitive and rhetorical demands. Rather than relying solely on policy dissemination or punitive measures, Saudi EFL higher-education contexts may benefit from instructional approaches that explicitly address areas of conceptual ambiguity and support students in developing nuanced, transferable understandings of academic integrity. Such approaches are essential not only for reducing unintentional plagiarism but also for fostering students' participation in academic discourse as competent and responsible writers.

A. Implications for Academic Writing Instruction

The findings of this study carry important implications for academic writing instruction in Saudi EFL higher education. First, the discrepancy between students' perceived confidence and their actual recognition of plagiarism suggests that instruction should move beyond general warnings and definitions. Instead, writing courses should prioritize explicit, scenario-based discussion of plagiarism, particularly in relation to paraphrasing, reuse of prior work, and AI-assisted writing. This approach aligns with research advocating the embedding of writing instruction within disciplinary contexts through genre-based and academic literacies approaches (Lea & Street, 2006; Wingate, 2012).

Second, the low recognition of self-plagiarism and teacher-provided material indicates a need for clearer instructional emphasis on authorship, ownership, and attribution as rhetorical choices rather than merely formal requirements. Incorporating targeted activities that require students to evaluate and revise borderline source-use cases may help bridge the gap between abstract knowledge and applied understanding. Such activities could provide the practice opportunities that research suggests are essential for successful transfer of paraphrasing knowledge (Liao & Tseng, 2010; Mahmud & Mohd Tahir, 2024).

Third, the findings support integrating plagiarism instruction into broader academic literacy development. Teaching citation and source use alongside paraphrasing and synthesis skills helps students understand plagiarism as part of effective academic communication. When students recognize that proper attribution is not merely about avoiding detection but about participating in scholarly conversation and knowledge construction, they are better positioned to engage responsibly with sources (Bazerman, 2013; Hyland, 2009).

Finally, given the emerging uncertainty surrounding AI-generated writing, instructors should proactively address how such tools can and cannot be used within academic contexts. Explicit guidance on ethical AI use may reduce confusion and support students in navigating evolving authorship norms. These recommendations collectively point toward educative rather than punitive institutional approaches, which research suggests are more effective in supporting students' development as responsible academic writers (Hu & Lei, 2015; Razi & Şahan, 2023).

B. Limitations and Recommendations for Future Research

Several limitations should be considered when interpreting the findings of this study. First, the data were collected from a single institution within a Saudi higher-education context and involved female undergraduates enrolled in Research Writing at King Khalid University. The all-female sample reflects the gender-segregated structure of the institution rather than a deliberate focus on gender differences in plagiarism understanding. While the findings are relevant to similar EFL academic writing contexts, they should not be generalized beyond comparable institutional and instructional settings. Future research could examine whether similar patterns of plagiarism recognition emerge among male Saudi EFL students.

Second, the study relied on self-reported questionnaire data and scenario-based judgments rather than direct analysis of students' written texts. Although scenario-based instruments are effective in revealing conceptual understanding, future research could triangulate these findings with textual analysis to examine how plagiarism understanding is enacted in actual writing practices, as demonstrated in the paraphrasing tasks used by Liao and Tseng (2010).

Third, while the questionnaire captured key areas of plagiarism recognition, the limited number of items (two for perceived understanding) constrains the depth of analysis possible for Research Question 1. Future studies should consider using more comprehensive measures of perceived plagiarism knowledge. Additionally, the study did not explore the instructional histories or prior educational experiences that may have shaped students' perceptions. Including qualitative data, such as interviews or reflective responses, could provide deeper insight into the reasoning behind students' judgments and the cultural or educational factors influencing their understanding.

Finally, the analysis focused on descriptive patterns rather than causal explanations. Future studies could investigate how targeted instructional interventions influence students' recognition of plagiarism over time, particularly in relation to paraphrasing and emerging forms of technology-assisted writing.

REFERENCES

- [1] AbdAlgane, M., Ali, R., Othman, K., Ibrahim, I. Z. A., Alhaj, M. K. M., Ezzeldin, M. T., & Ali, F. S. A. (2025). Exploring AI-generated texts vs. human-written texts in EFL academic writing: A case study of Qassim University in Saudi Arabia. *World Journal of English Language*, 16(2), 114–131. <https://doi.org/10.5430/wjel.v16n2p114>
- [2] Alhadlaq, A. S., Dahmash, A. B., & Alshomer, F. (2020). Plagiarism perceptions and attitudes among medical students in Saudi Arabia. *Sultan Qaboos University Medical Journal*, 20(1), e77–e82. <https://doi.org/10.18295/squmj.2020.20.01.011>
- [3] Alsaedi, N. S., & Alhums, M. H. (2024). Saudi undergraduate students' perceptions of plagiarism: A case of EFL research writing tasks during e-learning sessions. *Heliyon*, 10(22), Article e39804. <https://doi.org/10.1016/j.heliyon.2024.e39804>
- [4] Alsofyani, A. H., & Barzanji, A. M. (2025). The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: A mixed-methods study. *Journal of Educational Computing Research*, 63(2), 399–428. <https://doi.org/10.1177/07356331241307297>
- [5] Alua, M. A., Asiedu, N. K., & Bumbie-Chi, D. M. (2023). Students' perceptions of plagiarism and usage of Turnitin anti-plagiarism software: The role of the library. *Journal of Library Administration*, 63(1), 119–136. <https://doi.org/10.1080/01930826.2022.2148659>
- [6] Bazerman, C. (2013). *A rhetoric of literate action: Literate action volume I*. WAC Clearinghouse.
- [7] Chu, S. K. W., Hu, X., & Ng, J. (2020). Exploring secondary school students' self-perception and actual understanding of plagiarism. *Journal of Librarianship and Information Science*, 52(3), 806–817. <https://doi.org/10.1177/0961000619860726>
- [8] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- [9] Farahat, A. (2022). Elements of academic integrity in a cross-cultural middle eastern educational system: Saudi Arabia, Egypt, and Jordan case study. *International Journal for Educational Integrity*, 18, Article 9. <https://doi.org/10.1007/s40979-021-00095-5>
- [10] Howard, R. M. (1995). Plagiarisms, authorships, and the academic death penalty. *College English*, 57(7), 788–806. <https://doi.org/10.2307/378403>
- [11] Hu, G., & Lei, J. (2015). Chinese university students' perceptions of plagiarism. *Ethics & Behavior*, 25(3), 233–255. <https://doi.org/10.1080/10508422.2014.923313>
- [12] Hyland, K. (2009). *Academic discourse: English in a global context*. Continuum.
- [13] Issrani, R., Alduraywish, A., Prabhu, N., Alam, M. K., Basri, R., Aljohani, F. M., Alolait, M. A., Alghamdi, A. Y., Alfawzan, M. M., & Alruwili, A. H. (2021). Knowledge and attitude of Saudi students towards plagiarism—A cross-sectional survey study. *International Journal of Environmental Research and Public Health*, 18(23), Article 12303. <https://doi.org/10.3390/ijerph182312303>
- [14] Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, 15(4), 261–278. <https://doi.org/10.1016/j.jslw.2006.09.006>
- [15] Lea, M. R., & Street, B. V. (2006). The “academic literacies” model: Theory and applications. *Studies in Higher Education*, 31(2), 159–172. <https://doi.org/10.1080/03075070600572116>
- [16] Li, Y., Flowerdew, J., & Cargill, M. (2024). How English academic writing textbooks written by Chinese EFL teachers address the issue of plagiarism. *System*, 123, Article 103327. <https://doi.org/10.1016/j.system.2024.103327>
- [17] Liao, M.-T., & Tseng, C.-Y. (2010). Students' behaviors and views of paraphrasing and inappropriate textual borrowing in an EFL academic setting. *Journal of Pan-Pacific Association of Applied Linguistics*, 14(2), 187–211.
- [18] Liu, M., & Wu, Y. (2020). Chinese undergraduate EFL learners' perceptions of plagiarism and use of citations in course papers. *Cogent Education*, 7(1), Article 1855769. <https://doi.org/10.1080/2331186X.2020.1855769>
- [19] Macfarlane, B., Zhang, J., & Pun, A. (2014). Academic integrity: A review of the literature. *Studies in Higher Education*, 39(2), 339–358. <https://doi.org/10.1080/03075079.2012.709495>
- [20] Madkhali, M. M. (2017). *Saudi students' perception of plagiarism* [Master's thesis, St. Cloud State University]. Repository @ St. Cloud State. https://repository.stcloudstate.edu/engl_etds/91/
- [21] Maharajh, L. R. (2020). Exploring university students' perceptions of plagiarism: A focus group study. In *Proceedings of the International Conference on Multidisciplinary Research* (pp. 54–69). <https://doi.org/10.26803/MyRes.2020.05>
- [22] Mahmud, N., & Mohd Tahir, M. H. (2024). Mitigating plagiarism in ESL academic writing: Evaluating the efficacy of educational intervention. *Journal of Language and Education*, 10(3), 53–68. <https://doi.org/10.17323/jle.2024.21568>
- [23] Muluk, S., Habiburrahim, H., & Safrul, M. S. (2021). EFL students' perceptions of plagiarism in higher education: Triggering factors and avoiding strategies. *Jurnal Ilmiah Didaktika*, 22(1), 20–36. <https://doi.org/10.22373/jid.v22i1.9154>
- [24] Nketsiah, I., Imoro, O., & Barfi, K. A. (2024). Postgraduate students' perceptions of plagiarism, awareness, and use of Turnitin text-matching software. *Accountability in Research*, 31(7), 786–802. <https://doi.org/10.1080/08989621.2023.2239736>
- [25] Park, C. (2003). In other (people's) words: Plagiarism by university students—Literature and lessons. *Assessment & Evaluation in Higher Education*, 28(5), 471–488. <https://doi.org/10.1080/02602930301677>
- [26] Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12(4), 317–345. [https://doi.org/10.1016/S1060-3743\(03\)00042-0](https://doi.org/10.1016/S1060-3743(03)00042-0)
- [27] Pecorari, D. (2008). *Academic writing and plagiarism: A linguistic analysis*. Continuum.
- [28] Pecorari, D., & Petrić, B. (2014). Plagiarism in second-language writing. *Language Teaching*, 47(3), 269–302. <https://doi.org/10.1017/S0261444814000056>
- [29] Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing*, 21(2), 149–164. <https://doi.org/10.1016/j.jslw.2012.03.003>
- [30] Razi, S., & Şahan, Ö. (2023). From transgressors to authors: Promoting EFL writing through academic integrity integrated instruction. *System*, 117, Article 103108. <https://doi.org/10.1016/j.system.2023.103108>
- [31] Sheerah, H. A. H. (2024). Perceptions of digital academic dishonesty in English writing at applied colleges of King Khalid

- University. *World Journal of English Language*, 14(4), 620–630. <https://doi.org/10.5430/wjel.v14n4p620>
- [32] Shen, Y., & Hu, G. (2021). Chinese graduate students' perceptions of plagiarism: A mixed-methods study. *Accountability in Research*, 28(4), 197–225. <https://doi.org/10.1080/08989621.2020.1869386>
- [33] Silfiani, S., Aziz, Z. A., & Daud, B. (2018). Plagiarism in English academic writing: Students' perceptions. *English Education Journal*, 9(1), 102–117. <https://download.garuda.kemdikbud.go.id/article.php?article=1564641>
- [34] Tran, M. N., Hogg, L., & Marshall, S. (2022). Understanding postgraduate students' perceptions of plagiarism: A case study of Vietnamese and local students in New Zealand. *International Journal for Educational Integrity*, 18(1), Article 3. <https://doi.org/10.1007/s40979-021-00092-2>
- [35] Waltzer, T., & Dahl, A. (2023). Students' perceptions and evaluations of plagiarism: Effects of text and context. *Journal of Moral Education*, 52(1), 46–62. <https://doi.org/10.1080/03057240.2020.1779047>
- [36] Wingate, U. (2012). Using academic literacies and genre-based models for academic writing instruction. *Journal of English for Academic Purposes*, 11(1), 26–37. <https://doi.org/10.1016/j.jeap.2011.11.002>

Nawal I. Alhodithi received her PhD in Applied Linguistics and Composition from Indiana University of Pennsylvania, USA, in 2020. She earned her MA in Teaching English as a Second Language (TESL) from the University of Central Oklahoma in 2014 and her BA in English from King Khalid University in 2002.

She is currently an Assistant Professor and Vice-Chair of the English Department at the College of Languages and Translation, King Khalid University, Abha, Saudi Arabia. She previously served as Head of the English Enhancement Unit and Supervisor of the Counseling and Advising Unit at the college. Her previous publications include work on native and non-native English speaker dichotomies and the significance of L2 teachers' cultural and linguistic backgrounds in composition classes. Her current research interests include EFL pedagogy, academic writing, sociolinguistics, learner identity, sense of belonging in higher education, and undergraduate research experiences.

Dr. Alhodithi can be contacted at nalhdithi@kku.edu.sa; ORCID iD: <https://orcid.org/0000-0001-7602-6336>

The Impact of AI-Assisted Tools on Enhancing the Quality of Legal Translation Outputs: A Case Study of Jordanian Students

Alalddin A. Al-Tarawneh

Department of English Language & Literature, Zarqa University, Zarqa, Jordan

Mohammed Al-Badawi

Department of English Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Muath Algazo

Department of English Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Rula Ahmad Mahmoud Abu-Elrob

Department of English Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Bilal Alsharif

Department of English Language and Literature, Mutah University, Al-Karak, Jordan

Nader Albkower

Department of Translation, Mutah University, Jordan

Abstract—This study seeks to examine the effects of the use of Artificial Intelligence (AI)-assisted tools (ChatGPT and DeepL) on the quality of legal translation outputs by Jordanian undergraduates. As legal translation necessitates a high degree of specificity, context, and accuracy in the utilization of terminology, the following research will explore whether AI can help students accomplish this requirement. To achieve the research objective, a case study that combined methods was employed. A sample of 68 third- and fourth-year students from Zarqa University and Mutah University who specialize in translation was chosen to participate in the study. The participants were required to complete a pre-test and a post-test, each of which included legal texts (statutes, contracts, and court rulings). The following day's session involved utilizing AI. The quantitative analysis of the data collected from the exams and surveys indicated how terminology, grammar, context, and fluency improved translation quality. Qualitative analysis of surveys and interviews revealed participants' perceptions of their experiences and resources. The study's findings showed that using AI technologies to improve terminology, grammar, context, and fluency resulted in increased translation quality. Students also believed that the tools increased their confidence and efficiency; however, they also had concerns about overreliance on AI and the potential lack of legal content handling ability in the tools. Finally, this study revealed the pedagogical potential of using AI tools in legal translation training while pointing to the need for continuous human control and critical thinking.

Index Terms—AI-Assisted translation, legal translation, ChatGPT, DeepL

I. INTRODUCTION

The translation of legal documents is one of the most complex areas of translation. It differs from general translation by a very high degree of close connection between the text being translated and the legal system, its terminological units, and the interpretive tradition of both the source and target languages. The translator is required to be not only highly proficient in the style conventions and legal terminology of both languages but also aware of the differences between the legal systems (Herschensohn, 2022). Otherwise, legal translation errors can have very serious consequences, including contractual disputes, misinterpretation of laws, or even miscarriages of justice. For students whose knowledge of the subject and language is still limited, the translation of legal texts is extremely difficult from an academic and professional perspective.

The integration of AI into the world of language technology has affected the way translators operate as well. As opposed to the previous generation of machine translation methods, the statistical (stat-based) and rule-based (RB) approaches to machine translation have had less success in the legal context of translation. The most recent development has been in the NMT/NLP (natural language processing) method, which has facilitated the creation of more intricate machine-driven tools that have increased the accuracy and efficiency of the translation of text. These instruments augment the power of

AI-versatile translation tools already possessed (Amirdabbaghian, 2025). DeepL, Google Translate, and OpenAI's recently launched ChatGPT use context-sensitive algorithms that generate more accurate and semantically coherent translations. Moreover, other machine translation tools, such as SDL Trados and memoQ, can also be made human-readable. This is useful for nonfiction books.

In education, these instruments based on AI are already commonplace in the language and translation fields and serve as extra content that enables students to assess their knowledge, receive feedback, and revise their projects multiple times. For those who translate, these can be utilized as cognitive tools, but they can also be used to improve their translation abilities.

The expanding use of legal translation in Jordanian higher education institutions has necessitated the teaching of such a specialized discipline. Many students have difficulty with vocabulary, syntactic structures, and legal reasoning when producing high-quality translations. The limited availability of bilingual legal resources and the differences between Jordanian legal traditions (based on civil, Islamic, and customary laws) and the source language (English) also create some issues (Shryock, 1995). Moreover, despite the growing use of AI-based tools, the area is still under-researched in terms of how AI tools impact the quality of students' legal translations.

Considering these findings, it is worthwhile exploring how these tools can be leveraged in the teaching of legal translation to help students acquire the necessary skills for legal translation. More specifically, it is important to examine whether these tools contribute to the improvement of students' translation accuracy, the suitability of the translation to the context, and terminological correctness.

A. Problem Statement

It is particularly important to mention that Jordanian students enrolled in the study of translation or law often struggle with legal translation due to their lack of legal terminology or legal style, as well as the inability to comprehend the legal systems of the target or source language. Traditional classroom teaching is often inadequate for equipping students with the cognitive and linguistic tools required for legal translation. Consequently, it is necessary to investigate alternative or complementary tools that could bridge the competency gap. This is because the accuracy of translation is essential, especially in legal contexts where accuracy has institutional and legal consequences. Therefore, the use of AI-assisted tools is a potential opportunity that has not been explored to the extent it could be in helping student translators fulfill these needs. Nonetheless, it is unclear how these techniques contribute to improved translation quality in legal contexts, if at all.

B. Purpose of the Study

This study's purpose is to examine the effect of the application of Artificial Intelligence (AI) tools to the process of legal translation among university students in Jordan. The study will compare the outputs and attitudes of students to these technologies. It will also investigate whether these technologies have an impact on the accuracy and context of legal translation outputs.

C. Research Questions

1. How do AI tools affect the accuracy of legal translation by Jordanian students?
2. How do students evaluate the application of AI tools in legal translation?
3. Can the use of AI tools improve the quality of legal translation?

D. Significance of the Study

The results of this study will have an impact on the field of AI-assisted translation teaching, including legal translation teaching, by allowing curriculum developers, translation teachers, and students to recognize the educational and practical value of utilizing AI-assisted translation tools in the competitive field of law.

E. Scope and Limitations

This study focuses on students who underwent legal translation training as part of their undergraduate studies, whether in translation, applied linguistics, or English language and literature majors. The participants were recruited from two Jordanian universities: Zarqa University and Mutah University. The research specifically aims to assess how AI tools affect the quality of legal translation tasks that these students produce. The study has several limitations. First, it is restricted to undergraduate students whose level of translation proficiency may not be as adept as that of professional legal translators or graduate students. Second, the research is restricted to a limited number of legal texts used in the undergraduate legal translation courses. Third, it focuses on a limited set of AI tools, such as ChatGPT and DeepL, and not the full array of translation technologies available. Therefore, the results of this study can be understood within this specific educational and institutional context.

II. LITERATURE REVIEW

A. Legal Translation: Definitions and Characteristics

Legal translation refers to the translation of legal texts and documents. It also requires that the texts and documents retain their original meaning and purpose (Kaplow, 1994). Legal translation is extremely specific; any misconception or

ambiguity could lead to misunderstandings or legal issues. Legal translation is distinguished by its specialized nature, the use of specialized terminology and structures, and the necessity of considering the cultural and legal context of the source text to achieve the desired effect in another jurisdiction (Evans et al., 2015). The main characteristics of legal translation include:

1. Legal translators must ensure accuracy; an error or omission could render an agreement invalid or cause a dispute (Simester, 1995). Law is precise, and legal language must be formal.
2. Legal documents are full of terms of art and jargon with no clear counterparts in other languages. Translators have to be aware of the precise legal meaning of such terms and sometimes even produce innovative solutions to express them. Studies note that legal translators have to analyze the source's rhetorical and syntactic structures to capture its implicit legal implications (Mediouni, 2016). In practice, this means having extensive bilingual legal lexicons and knowing how, or if, to domesticate concepts for the target legal system.
3. The law is always closely related to the culture. That is, the meanings of many legal terms, concepts, and provisions are based on the social and legal culture of the source country, so the translator needs to understand the culture behind the text to translate it in a way that will be valid in the target culture (Alcaraz, 2017). For example, the term "interest" as it relates to property or family relationships may have no corresponding concept in another culture, which requires an explanatory translation or adaptation. A translator should also consider the legal norms, the legal system, and the values of the society in the target country to ensure that the translated document is not only linguistically correct but also legally equivalent and socially appropriate in the target jurisdiction (Prieto-Ramos, 2017).

In short, legal translation is not just the domain of the translator but also of a lawyer, combining linguistic, subject-matter, and cultural knowledge (Al-Tarawneh, 2024). Legal translation is sometimes described as the task of a comparative lawyer. Such a translator should compare and contrast two legal systems and should be able to "mediate" between them. Accuracy of wording and the faithfulness of the legal intent are of utmost importance. This is because translated contracts, statutes, or court decisions are as binding as the original ones. Professional legal translators, therefore, are highly sensitive to the fact that their translation is supposed to be subjected to strict scrutiny and has to be correct and conform to the conventions of the target legal system (Procházka, 2022).

B. AI in Translation: A General Overview

In the last decade, artificial intelligence has advanced rapidly, and this includes in the field of translation. Rule-based and statistical machine translation methods have been replaced by Neural Machine Translation (NMT), which utilizes deep neural networks to learn linguistic patterns (Laurençon et al., 2022). The ability of modern AI translation tools to produce fluent and coherent translations at lightning speed is due to deep learning algorithms that are trained on vast bilingual corpora (Laurençon et al., 2022). By utilizing the most recent Natural Language Processing (NLP) methods, machine translation now has a greater capacity to determine context and subtlety, rendering the older methods of translation ineffective. This suggests that AI-driven machine translation produces more accurate and context-dependent results than previous methods and may be employed in real time in everyday contexts (Cheng et al., 2023).

Translation is an area of human activity in which technological change has contributed enormously to global communication (Shubham, 2024). A new term has been coined to refer to this type of translation—"AI translation." It is based on the premise that machine translation is carried out by humans (Selincourt, 2016). The most well-known examples of this technology are the online translation tools Google Translate and Microsoft's Bing Translate. They provide instant translations to millions of users with quality sufficient to meet their needs. Even though translations, especially in some more specialized fields, are often inadequate, users still choose them because they are fast and easy to use. The translation industry has almost completely integrated computer-aided translation (CAT) tools, providing translators with glossaries, translation memories, and other interfaces. This suggests that artificial intelligence can also support translators' workflows.

The rapid development of machine learning has created major breakthroughs in the field of machine translation, and large-scale neural language models have emerged rapidly as a result. In late 2022, OpenAI's ChatGPT model (GPT-3.5) proved that large-scale language models can be used in fields such as translation (Tian & Deng, 2024). Since ChatGPT and its successors are trained on large amounts of text and fine-tuned with more advanced learning techniques, the quality of AI-assisted translation has been significantly improved. In 2023, the multimodal model GPT-4 was released, which can accept text and images as input. GPT-4 has achieved near-human performance in various language tests, including translation. Although these models are not designed specifically for translation, their large scale and dialogue-based training enable them to analyze complex sentences and process long text paragraphs. Although the most effective AI models still cannot completely avoid errors (some common and specialized texts are still difficult to understand), the gap between machine translation and human translation has been significantly narrowed thanks to artificial intelligence technology (Acquaviva et al., 2022). This summary points to a vital pattern: artificial intelligence went from being a luxury to a necessity, and work on neural networks and natural language processing underpinned its relentless advance.

C. AI-Assisted Tools Commonly Used in Translation (ChatGPT)

ChatGPT is a powerful and versatile AI translation tool. It is a large language model (LLM) that can be fine-tuned to achieve conversational responses and perform a variety of language tasks (such as translation) even if it is not specifically

trained for a particular domain. ChatGPT is based on the GPT family of transformer architectures and can understand context and generate high-quality multilingual text using its multilingual training data. In specific tests, researchers demonstrated that ChatGPT possesses the ability to fluently handle contextual details and domain specificity, on par with domain-trained machine translation systems (Howard, 2013). For example, in one of these specific evaluations, ChatGPT outperformed leading neural machine translation (DeepL) systems by intuitively understanding the implicit meaning in literary translations and producing more concise, natural, and nuanced translations.

ChatGPT is used as a translation tool in professional settings. Its chat-like user interface allows translators to repeatedly edit translations. For example, a user can enter a sentence or paragraph and request that it be rephrased, rewritten with a different tone, or that different terms be used in subsequent queries. This interactive capability is considered an improvement over the one-off results of machine translation and enables basic collaboration between humans and artificial intelligence. Several professional studies have shown that translators find ChatGPT useful for generating a preliminary translation and proposing solutions for complex segments that could be edited to improve accuracy (Xu & Tan, 2025). Creating a draft from the source text (as opposed to translating the source text) frees up experts' time to review and proofread the text. Furthermore, because ChatGPT's native language can be used for proofreading/editing, translators have been observed using ChatGPT to paraphrase/correct translations, using their native language to improve the readability of their work (Choi, 2005). Another observed use case is ChatGPT explaining difficult terminology/concepts, etc. For example, a student or translator might ask ChatGPT to explain a term or provide an example of its use in the language to better understand it before translating the term.

Conversely, ChatGPT also has some limitations. While it can generate fluent translations, these can be inaccurate in terms of facts or terminology. For example, because the model is unfamiliar with legislation and lacks access to the most up-to-date legal databases, it might use incorrect legal terms or overlook certain subtleties. A recent survey of translators' opinions on ChatGPT revealed concerns about the system's reliability and competence with sensitive content (Wu et al., 2024). When working with sensitive texts, concerns may arise about the possibility of machine translations containing factual or cultural errors or about the confidentiality of client documents submitted to these systems. For example, in legal translation, ChatGPT may correctly translate a contractual clause, but without adopting the normal legal terminology or style of the target legal system (Haber et al., 2025). This means that the translated text will not be valid for official purposes.

Taken together, these considerations reinforce the view that ChatGPT is a useful tool, but it is not a substitute for human expertise. Best practice in the field is to use ChatGPT as an aid to accelerate and assist translation tasks (for example, for routine text or alternative phrasing), but a human translator should ensure the accuracy of legal terms and take responsibility for critical review and approval. In short, ChatGPT is an advanced translation technology powered by artificial intelligence. Its speed is faster than ever. Its fluidity is incredible. However, for it to be applicable in the professional legal industry, it still requires human control, wisdom, and the moral standards of its users.

D. Previous Research on AI in Legal Translation

With the development and application of AI tools like ChatGPT, the amount of research dedicated to evaluating AI performance in legal translation has increased in recent years. This type of research typically focuses on comparing human and machine translation. Studies before and after 2020 have shown that, while AI systems have made impressive progress in their performance, human translation still outperforms machine translation overall, particularly in terms of terminology and pragmatics. For example, Al-Romany, Tahseen, and Kadhim found a significant difference in quality between AI-generated and human-generated legal translation (Al-Romany et al., 2024). They found that human translation outperformed AI in capturing the distinctive features of legal language. They suggested that these features were formality, legal meaning, and context, which are easier for humans to understand (Al-Romany et al., 2024). Experience and expertise were found to benefit, as professional translators were able to correctly understand and reproduce complex legal expressions where AI was unable to do so (Al-Romany et al., 2024). In line with this, another study found that AI translations of legal texts often misinterpret legal terms or use them incorrectly, leading to ambiguities that a human expert would not make (Alqahtani, 2025). This ambiguity can be crucial in legal contexts, as mistranslating a term can impact the outcome of a case, so automatic legal translation is viewed with caution.

Many studies have since focused on comparing the performance of ChatGPT with other translation methods. Ding (2024) compared the performance of ChatGPT with four commercial NMT systems on the task of translating legal English to Chinese (Ding, 2024). Results were mixed and dependent on the translation direction: in the English-to-Chinese direction, both ChatGPT and the NMT engines performed well, with ChatGPT's translation quality being slightly lower but not statistically significantly different from the specialized NMT systems (Ding, 2024). This means that, for some language pairs and directions, ChatGPT can achieve performance comparable to dedicated machine translators. However, in the English-to-Chinese legal translation direction, neither ChatGPT nor the machine counterparts achieved an acceptable accuracy level. On average, NMT systems beat ChatGPT in this area, and more importantly, ChatGPT generated more errors, which were often more significant in nature. Ding noted that the types of errors made by ChatGPT and the NMT systems were similar (e.g., wrong legal terms, incorrect translations of complex structures), but ChatGPT made those errors more frequently (Ding, 2024). These results from the Chinese context support the hypothesis that ChatGPT's translation performance can vary; it can do well in one language direction but fail in another, perhaps due to differences in the coverage of training data or the complexity of the source text structures. They also reinforce the

assumption that legal translation presents a particularly challenging test for artificial intelligence, given the importance of accuracy and the differences between legal systems.

Zayed (2024) analyzed the English-Arabic translations of legal documents generated by ChatGPT (GPT-4 model) and compared them with human translations using various indicators, including the accurate use of legal terminology, the ability to translate complex and detailed sentences from the original text into easily understandable Arabic, and the degree of conformity with the norms of the local legal system. The main finding of this research was that human translation significantly outperformed the AI-generated results across all indicators (Zayed, 2024). In particular, human translators demonstrated a greater understanding of legal terminology and were able to translate complex legal structures more effectively, maintaining the formal style and typical expressions of local Arabic present in the Jordanian legal documents. In contrast, ChatGPT translations were less fluent and lacked context, failing to adequately reflect the use of technical terms and complex sentences in the legal documents (Zayed, 2024). For example, the author states that ChatGPT uses many words that can be characterized as direct transliterations of English legal terms into Arabic, albeit without full identity with the corresponding Arabic word. In other, albeit rare, cases, ChatGPT omits legal aspects (such as modal verbs or conditional clauses) that are essential for formulating legally binding documents such as contracts or regulations. As a result, the AI-generated text does not have the force of law and cannot be used as a legal document. The author concludes that AI tools such as ChatGPT or its equivalents have a very bright future (especially as they develop), but human translators are still essential, especially in legal texts (Zayed, 2024). Moreover, Zayed (2024) recommends harnessing the best of both worlds approach where, for example, AI is used to assist humans by speeding up first drafts or providing suggestions, but the final translations are edited and vetted by human experts. This is in line with the consensus in the literature that AI can be useful for improving the efficiency of legal translation, but accuracy, liability, and a more profound understanding of law mean that AI is best used as an assistive technology rather than as a stand-alone solution (Zayed, 2024).

To conclude, it should also be noted that the interest in AI for legal translation is international and increasing. Outside the Middle East and East Asia, researchers in Europe and other regions have also studied this topic. For example, Moneus and Sahari (2024) conducted a contrastive study of legal text translations in several languages and found similar results: that human translation outperforms AI on important legal fidelity metrics despite the performance gain in speed (Moneus & Yousef, 2024). Across studies, a recurring theme is that legal translation is a “stress test” for AI systems because of the need for complete clarity and correctness, and thus, one of the last domains where human expertise outperforms automation (Moneus & Yousef, 2024). However, technical improvements continue to come each year, and researchers refine evaluation methods (both automated metrics and human assessments) to keep track of the narrowing gap. In the literature thus far written on the subject, a cautious perspective emerges that AI-assisted legal translation is useful and improving, but human oversight is necessary to ensure terminological precision, legal validity, and all the subtle meanings that legal texts entail.

E. The Jordanian Context

Jordan's legal and educational climate specifically allows for the use of AI in the legal translation of texts. Though the Arabic language is the official language of Jordan, English is also present in contracts and other formal documents of law, in business law, and in academic texts. As such, high-quality Arabic-to-English and English-to-Arabic translations are crucial services. Alzghoul et al. (2025) argue that legal translation is a specialty in the country. Jordanian universities have incorporated this into their educational program. For example, the University of Jordan's master's degree in translation offers a class dedicated to legal translation. During the course, students confront the difficulty of the English language through learning legal terminology, the recognition of the conventions of legal phrases, and the importance of language in the legal sphere. They are responsible for converting legal documents (contracts, statutes, court judgments) from Arabic to English and the other way around, with an emphasis on both clarity and precision. The description of the course is an example of what legal translation certification is like for Jordan's courts and other professional organizations. Documents must be translated and proofread by qualified translators. Some translators hold official certifications (Welchman, 2000). If an English contract is to be registered as an Arabic translation, it must be legally binding in court. Thus, there are also certified translators for this. Translation errors or minor inaccuracies can have legal implications, so their training is also quite comprehensive. Jordan's legal system is based on civil law, but it is also influenced by Islamic law; as a result, translators also need to know comparative law (Welchman, 2000). All these factors make Jordan both in need of high-quality legal translation and acutely aware of the obstacles involved.

The integration of AI into various educational contexts has long been a matter of concern, and AI-based writing tools such as ChatGPT have begun to pose questions in both the classroom and in professional practice (Alazzam et al., 2024). One recent investigation has examined the introduction of ChatGPT in Jordanian higher education, and its findings are consistent with those in other countries. Students in Jordan appear to be open to using ChatGPT in their studies, and their desire to use the technology is correlated with their perception of ease of use and usefulness in their studies (Ajlouni et al., 2023). That is, students will embrace the technology when they see it as saving them time or helping them learn the material (and it is easy to access and use), despite the potential pitfalls. The results of a study conducted in Jordan indicate that students are aware of the disadvantages of AI-based tools. However, misinformation or data privacy issues were not substantial factors in the overall attitude toward ChatGPT (Altikriti & Ziad, 2025). This suggests that so long as its utility is clear, students are less likely to let worries about misuse or other negatives get in the way. For instructors, conversations

have begun about integrating AI into curricula, with experts arguing that tomorrow's translators should be equipped to use AI tools and to post-edit machine output. For example, a regional study emphasizes the importance of integrating machine translation and post-editing training into translator training programs as part of digital literacy development (Guerberof & Joss, 2019). At the more practical level, some faculty in Jordan are probably already letting students use ChatGPT or similar tools to write their translations and then teaching them to review and revise the output.

Legal translation is a highly technical field that demands a level of accuracy that often surpasses general translation, and Jordanian students and professionals are beginning to explore the use of AI-assisted tools to meet these demands. Indeed, ChatGPT has achieved enough fame in Jordan that legal practitioners have begun to recognize it as a new tool that could have potential in legal translation (Beirat et al., 2025). There are some reports that law students have been using ChatGPT to understand complicated cases or even draft bilingual moot court materials. However, because of the seriousness of legal translations, university administrators and instructors remind students not to rely on AI-produced outputs for legal translations. Rather, the narrative is that ChatGPT can support learning, as it can provide a fast draft or even explain a legal concept, and the student can verify and refine the AI output. That is, the benefits (speed, exposure to more fluent language patterns) are maximized, and the risks (students relying on the AI) are mitigated through human oversight. Jordanian policymakers and academics have begun to discuss policies for AI use, such as ensuring that, when students use online AI tools, they respect the ethical and privacy aspects and that AI does not hinder the development of students' translation skills. Generally, the response to AI has been optimistic but cautious. With the proper training, AI tools can aid in the quality of legal translations by Jordanian students, as long as the students are taught to assess and correct the AI's recommendations (Emara, 2024). The case study discussed here, which focused on Jordanian students, provides data on how ChatGPT might support (or in what ways it falls short in) legal translations produced by students. In the end, Jordan's experience may be representative of the rest of the world—embracing technology in translation education but in a way that maintains the accuracy and cultural appropriateness that have defined Jordanian legal translation for centuries.

III. METHODOLOGY

A. Research Design

This research was conducted using a case study design with a mixed-methods approach (quantitative and qualitative) to analyze the impact of using AI-assisted tools (ChatGPT and DeepL) on the quality of translated texts produced by Jordanian university students. To assess the accuracy and quality of translated texts before and after using AI-assisted tools, the researchers collected quantitative data. Qualitative data were also used to understand students' experiences, opinions, and attitudes towards using AI-assisted tools in translation work.

B. Participants

The sample consisted of 68 undergraduate students, 35 females and 33 males, studying translation, applied linguistics, or English language and literature. The participants were either juniors or seniors and had at least one course in legal translation. The participants were recruited from both Zarqa University and Mutah University, two Jordanian universities that specialize in training in translation and linguistics. All participants were selected on the basis of their basic familiarity with the legal translation tasks and terminology. Tables 1 and 2 present the distribution of the participants.

TABLE 1
DISTRIBUTION OF PARTICIPANTS BY UNIVERSITY AND GENDER

University	Female	Male	Total
Zarqa University	21	19	40
Mutah University	14	14	28
Total	35	33	68

TABLE 2
DISTRIBUTION OF PARTICIPANTS BY UNIVERSITY AND ACADEMIC YEAR

University	3rd Year (Juniors)	4th Year (Seniors)	Total
Zarqa University	22	18	40
Mutah University	13	15	28
Total	35	33	68

C. Instruments and Materials

The following instruments and materials were used in the study:

1. Pre- and Post-Tests on Legal Translation: Students completed two translation tasks, one before and one after they had the opportunity to use the AI tools. The texts were chosen to represent a range of legal genres, including contracts, court rulings, and statutes.
2. AI-Assisted Tools: During the follow-up test, students had access to ChatGPT and DeepL as artificial intelligence tools that they could use to create or improve their translations.
3. Survey Questionnaire: Students completed a questionnaire that included Likert scale questions about their experiences using AI-assisted tools. These questions covered the tools' perceived usefulness, perceived ease of

use, trust in the results, and willingness to use the tools in the future.

D. Data Collection Procedures

The process of collecting data occurred in these key stages over four weeks:

1. Orientation and Pre-Test: The research objectives and procedures were explained to the participants. They then conducted a pilot translation test (without using AI) by translating a random excerpt from a legal text (sample document provided) into English.
2. AI Tools Introduction: Students were trained in using ChatGPT and DeepL for legal translation, providing effective support, and analyzing the results.
3. Post-Test (with AI tools): At the end of the test, students completed a translation task (of the same difficulty and scope as previous tasks) and had access to AI tools. After completing the task, participants uploaded their final AI-assisted translations.
4. Interviews: After the following assessments, all participants completed a form. Semi-structured interviews were conducted with some participants to learn more about their experiences and ideas.

E. Data Analysis Methods

(a). Quantitative Analysis

The translation outputs from both the pre-test and post-test were evaluated using a standardized evaluation rubric that considered the following aspects:

1. Terminological accuracy.
2. Grammatical and syntactic structure.
3. Contextual and legal appropriateness.
4. Fluency and coherence.

Two human raters, both with legal translation experience, independently evaluated each translation to ensure that the two raters had a high degree of consistency. Statistical tests (e.g., paired t-tests) were conducted to assess whether the quality of the translation increased following the use of the AI platform.

(b). Qualitative Analysis

Descriptive statistics were employed to describe the results of the survey and the general attitudes and perceptions of the participants. Thematic analysis was employed to explore the primary themes regarding the students' perception of:

1. AI tool usefulness and limitations.
2. Confidence in AI-generated content.
3. Ethical concerns (e.g., overreliance, data privacy).
4. Personal strategies for integrating AI assistance into their workflow.

Data were interpreted using the six-stage model of Braun and Clarke to ensure the appropriate interpretation of the qualitative data (Braun & Victoria, 2006).

IV. RESULTS

A. Quantitative Findings

The effect of the AI-based tools on the quality of legal translation was determined by comparing the students' work prior to and following the translation exercises. To accomplish this, rubrics were created to assess the four areas: the accuracy of terminology, the syntax of speech, the context of the speech, and its fluency. Students could earn a maximum of 40 points for each translation task, with a maximum of 10 points for each of the four aspects mentioned above.

TABLE 3
OVERALL IMPROVEMENT

Measure	Pre-Test Mean Score	Post-Test Mean Score	Mean Gain
Terminology Accuracy	6.2	8.1	+1.9
Syntactic Correctness	6.7	8.4	+1.7
Contextual Appropriateness	5.9	7.8	+1.9
Fluency and Coherence	6.5	8.6	+2.1
Total Score (Out of 40)	25.3	32.9	+7.6

Paired-samples t-test results show that the overall translation quality was significantly improved after using ChatGPT and DeepL.

B. Qualitative Findings

(a). Student Perceptions

Surveys and interviews show that most students believe AI tools are very helpful in the translation process. Key points include:

1. Usability: 82% of students found ChatGPT and DeepL simple to utilize.

2. Confidence: 74% of students believed that AI could improve the translation process.
3. Awareness: 65% of respondents knew that AI-generated translations should be reviewed before submission.

(b). Benefits Identified

Students frequently cited the following benefits:

1. Terminology Inquiry: Artificial intelligence tools helped students understand legal terminology they previously struggled to remember.
2. Grammar and Style: Because of ChatGPT's recommendations, their sentence accuracy and length increased.
3. Speed: Time is saved in completing task drafts.

(c). Concerns Raised

Despite the positive feedback, students expressed several concerns:

1. Overdependence: Some students expressed concern about relying too heavily on AI without training themselves.
2. Accuracy: Some students were concerned that they were entrusting too many tasks to the AI without developing their own skills.
3. Context Sensitivity: 40% of students said that artificial intelligence sometimes fails to understand context.

C. Emerging Patterns and Themes

The following outcomes are exhibited after conducting a thematic analysis of the subject's content during the interview:

1. Terminology: Participating students broadened their understanding of legal terms and terminology. ChatGPT provided them with technical terms that they could practice or confirm.
2. Constant and Fluency: Consistency and fluency: The translations in the retest were more consistent in style, and the variations in tone or register were smaller. AI assistance had a longer consistency.
3. Contextual Understanding: The students increased their understanding of the different jurisdictions through the mistakes they observed regarding AI. This enhanced their understanding of the legal systems beneath them.
4. Workflow: With the help of AI, students were able to transform more swiftly and with less initial fear, with extra time devoted to reflection and revision. Many students described a process of first writing, then reviewing and improving it, using ChatGPT as a starting point and having their training to ensure accuracy.

V. DISCUSSION

A. Interpretation of Results

The results of this investigation demonstrated that the AI-versed tools (ChatGPT and DeepL) have a beneficial and quantifiable impact on the quality of legal translation that is performed by Jordanian students of law. The result of the post-test demonstrated a significant difference in all of the criteria (termed as terminology, syntactic accuracy, context-appropriateness, and fluency). This indicates that the tools have a significant impact on both linguistic and stylistic correctness in an academic setting.

Students found the suggestion of legal terminology and grammatical structures particularly helpful, as it avoided the errors and unclear wording that are often found in machine translation. However, the paper also reveals that AI is fallible. Students recognized that the jurisdiction-specific language and over-literal translations of legal statements were incorrect, demonstrating the limitations of AI in the specific knowledge and judgment of the domain, particularly regarding critical legal texts.

The results indicate that AI can enhance the quality of translation, but its effectiveness is dependent on the user's critical reading skills and legal translation training.

B. Implications for Translation Pedagogy

The implementation of AI-based technology in the translation field (primarily in the legal sphere) is a field that is plentiful with opportunities. The researchers believe that ChatGPT-like tools can be considered cognitive aids for students that help them to initially create a rough draft and explore alternative wording or terminology that they would not have considered.

In the legal field of translation, the incorporation of AI into teaching and learning is particularly significant:

1. Educate students on the importance of iteration; they will analyze the output of machines and revise it.
2. Help with developing post-editing abilities, which are of increasing importance in modern translation processes.
3. Support student reflection as students contrast their own translations with the suggestions provided by computers.

However, for AI to have a teaching value, instructors must be sure to explain the necessity of critical thinking and revision. To make sure the tools are employed effectively, students must be trained in the responsible use of AI tools (e.g., how to recognize errors and improve results to be legal).

C. Challenges and Ethical Considerations

Despite the educational value of these tools, they also raise several practical and ethical issues:

1. Additionally, students may rely on the output of AI and lose their passion for developing fundamental skills in legal translation and comparative law; this would diminish their potential for professional development in the

- long term.
2. Translations generated by AI are likely to be inaccurate, and small errors can have significant consequences in legal situations. Students may not realize this danger and may provide translations that are inaccurate or misrepresented without supervision.
 3. Students may participate in actual legal agreements or delicate legal texts via a chatbot like ChatGPT or DeepL, which poses questions about the security of data and the rights to intellectual property. Institutions must create formal rules that regulate the handling of data and intellectual property. The issues listed above are reasons for creating formal rules in academic environments that regulate the ethical behavior of AI, ensure human oversight, and define the acceptable methods of the classroom.

D. Comparison With Prior Research

The results of this study are in line with previous studies, both internationally and regionally. Like Ding (2024) and Zayed (2024), this study has found that ChatGPT and DeepL enhance fluency and lexical accuracy but perform poorly in the legal accuracy dimension, especially in context-bound and complex contexts. Also, the results of this study align with AI-Romany et al. (2024) in that the use of AI in legal translation requires human post-editing. The uniqueness of this study lies in the participants' group, as the focus is on Jordanian undergraduate students, a group that has not been adequately represented in the literature. The investigation, which focuses on the engagement of Jordanian undergraduate students with AI in a formal educational environment, provides information about the pedagogical benefits and challenges of AI in the legal translation field within the Jordanian context. In this regard, the results indicate that the utilization of AI in a guided and ethical manner can have a positive effect on student learning.

VI. CONCLUSION

The objective of this paper was to investigate the potential for AI-powered translation tools to enhance the quality of legal translations produced by Jordan's undergraduate students. Methods included pre- and post-tests, surveys, and interviews. The researchers conclude that the quality of the translations increased significantly after students had access to automated translation tools. More specifically, they significantly enhanced their grasp of terminology, grammar, context, and fluency.

Qualitative results demonstrated that the students found the AI tools beneficial, time-saving, and educational during the writing or revision phase of translations. On the other hand, students recognized the lack of automation in the management of legal terminology specific to a jurisdiction, the interpretation of complex legal concepts, and the verification of domain-specific accuracy.

Indubitably, this study enhances the existing knowledge of how to incorporate AI into translation instruction, specifically in the legal field and in the Arab-English context. It delves further into the Jordanian students' opinions and attitudes towards AI technology, specifically regarding ChatGPT, in the unexplored area of AI-assisted translation research. Additionally, it provides empirical evidence to support the pedagogical goal of AI-based tools in enhancing students' learning and efficiency while also facilitating the use of AI as a cognitive aid for legal translation. The pedagogical approach promises that AI should not be seen as a replacement for human translation but rather as an assistive tool.

A. Recommendations

Based on these findings, the study recommends the following: first, AI tools must be used as tools to enhance learning in legal translation courses, for example, by emphasizing the importance of critically reviewing the output and post-editing skills and by ensuring students understand the ethical implications of AI. Second, tasks that include comparing human and AI translations should be designed to allow students to reflect on the output and develop a better sense of their judgment. Third, clear guidelines should be developed for institutional usage of AI in courses (i.e., data privacy, academic integrity). Fourth, continuing education opportunities should be promoted for instructors to teach with and about AI tools. Fifth, domain-specific modules or legal translation plugins for improved content generation in specific areas should be developed. Finally, to allow users to choose legal systems or document types to help improve context, jurisdiction-aware capabilities should be developed.

B. Suggestions for Future Research

This case study opens several avenues for further investigation, the first of which is by researching more general demographic studies with postgraduate students, professional translators, and legal professionals to understand how different user groups interact with AI in legal translation. Next, there should also be tool-specific performance testing of different AI models (ChatGPT, DeepL, Google Translate) in different legal domains and languages. Additionally, other research should include longitudinal studies to understand how long-term use of AI in translation influences skill acquisition, accuracy, and reliance on the technology over time. Future research in this area will deepen our understanding of the evolving role of artificial intelligence in legal translation and help ensure that its integration into legal translation is both educational and ethical.

REFERENCES

- [1] Acquaviva, Sam, et al. (2022). Communicating natural programs to humans and machines. *Advances in Neural Information Processing Systems*, 35, 3731–3743.
- [2] Ajlouni, Aseel Omar, Fatima Abd-Alkareem Wahba, & Abdallah Salem Almahaireh. (2023). Students' attitudes towards using ChatGPT as a learning tool: The case of the University of Jordan. *International Journal of Interactive Mobile Technologies*, 17(18), 22–36.
- [3] Alazzam, T. S., Alzghoul, M. A., & Alzghoul, R. (2024). The translation of medical and health-related idioms by university students from English into Arabic: Challenges and strategies. *Jordan Journal of Modern Languages & Literatures*, 16(4), 941–954.
- [4] Alcaraz, Enrique, & Brian Hughes. (2014). *Legal Translation Explained*. London: Routledge.
- [5] Alqahtani, Demah Aamer. (2024). Transcending ambiguities: Enhancing AI-driven Arabic to English translations with human expertise. *Journal of Languages and Translation*, 11(3), 59–81.
- [6] Al-Romany, Hussein, Tahseen Ali, & Maryam Jawad Kadhim. (2024). Artificial intelligence impact on human translation: Legal texts as a case study. *International Journal of Linguistics, Literature & Translation*, 7(5), 1–14.
- [7] Altikriti, Sahar, & Ziad Nemrawi. (2025). The use of ChatGPT as a learning tool among Jordanian students in technical universities: Attitudes and perceptions. *International Journal of Interactive Mobile Technologies*, 19(5), 101–115.
- [8] Alzghoul, M. A., Abualasal, J., Alazzam, T. S., Al-Ghammaz, S. A. D., & Alzghoul, R. M. (2025). The status quo of translation technology tools in translator training programs at Jordanian universities. *World Journal of English Language*, 15(7), 398–398.
- [9] Al-Tarawneh, A. (2024). The impact of legal translation on criminal proceedings. *Pakistan Journal of Criminology*, 16(3), 793–804.
- [10] Amirdabbaghian, Amin. (2025). Latest developments in translation and interpreting technologies in Malaysia. In Mahmoud Ibrahim & Ahmad Als Salman (Eds.), *Cutting-Edge Approaches in Translator Education and Pedagogy* (pp. 191–216). Hershey, PA: IGI Global.
- [11] Beirat, Mohammad Amin, Dima Maher Tashtoush, Mohammad Ahmad Khasawneh, Eyad Abdulrahman Az-Zo'bi, & Mohammad Ahmad Tashtoush. (2025). The effect of artificial intelligence on enhancing education quality and reduce the levels of future anxiety among Jordanian teachers. *Applied Mathematics*, 19(2), 279–290.
- [12] Braun, Virginia, & Victoria Clarke. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- [13] Cheng, Yuxin, Zihan Zhang, Dandan Han, & Yang Liu. (2023). Context-based AI translation from a globalization perspective: A case study of ChatGPT. *Sino-US English Teaching*, 20(9), 370–380.
- [14] Choi, Jungwha. (2005). The role of translators and issues of choice: Faithfulness and readability in literary translations. *FORUM Revue internationale d'interprétation et de traduction*, 3(2), 99–116.
- [15] Ding, Lijie. (2024). A comparative study on the quality of English-Chinese translation of legal texts between ChatGPT and neural machine translation systems. *Theory and Practice in Language Studies*, 14(9), 2823–2833.
- [16] Emara, Eman Abd El-Hafeez Mohamad. (2024). Using AI tools to enhance translation skills among basic education English major students. *CDELT Occasional Papers in the Development of English Education*, 86(1), 339–380.
- [17] Evans, Lisa, Rachel Baskerville, & Katariina Nara. (2015). Colliding worlds: Issues relating to language translation in accounting and some lessons from other disciplines. *Abacus*, 51(1), 1–36.
- [18] Guerberof Arenas, Ana, & Joss Moorkens. (2019). Machine translation and post-editing training as part of a master's programme. *Jostrans: The Journal of Specialised Translation*, 31, 217–238.
- [19] Haber, Eldar, Lior Zalmanson, Einat Albin, & Roy Peled. (2025). *Using AI in Academic Writing and Research: A Complete Guide to Effective and Ethical Academic AI*. Cham: Springer Nature.
- [20] Herschensohn, Julia. (2022). *Bilingualism, Language Development and Processing Across the Lifespan*. Cambridge: Cambridge University Press.
- [21] Howard, Veena R. (2013). *Gandhi's Ascetic Activism: Renunciation and Social Action*. Albany: SUNY Press.
- [22] Kaplow, Louis. (1994). The value of accuracy in adjudication: An economic analysis. *The Journal of Legal Studies*, 23(S1), 307–401.
- [23] Laurençon, Hugo, et al. (2022). The BigScience ROOTS corpus: A 1.6 TB composite multilingual dataset. *Advances in Neural Information Processing Systems*, 35, 31809–31826.
- [24] Mediouni, Mohammed. (2016). Towards a functional approach to Arabic-English legal translation: The role of comparable/parallel texts. In Nadia Abdelwahab (Ed.), *New Insights into Arabic Translation and Interpreting* (pp. 115–160). Cham: Springer.
- [25] Moneus, Ahmed Mohammed, & Yousef Sahari. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10(6), 1–10.
- [26] Prieto-Ramos, Fernando. (2017). Global law as translated text: Mapping institutional legal translation. *Tilburg Law Review*, 22(1–2), 185–214.
- [27] Procházka, Radoslav. (2002). *Mission Accomplished: On Founding Constitutional Adjudication in Central Europe*. Bratislava: Slovak Academic Press.
- [28] Selincourt, Chris de. (2016). *Where Is the Mind of the Media Editor? An Analysis of Editors as Intermediaries Between Technology and the Cinematic Experience* [PhD dissertation]. Cardiff Metropolitan University.
- [29] Shryock, Andrew J. (1995). Popular genealogical nationalism: History writing and identity among the Balga tribes of Jordan. *Comparative Studies in Society and History*, 37(2), 325–357.
- [30] Shubham, M. (2024). Breaking language barriers: Advancements in machine translation for enhanced cross-lingual information retrieval. *Journal of Electrical Systems*, 20(9s), 2860–2875.
- [31] Simester, Andrew P. (1995). Why omissions are special. *Legal Theory*, 1(3), 311–335.
- [32] Tian, Chuanmao, & Deng Juntao. (Eds.). (2024). *Translating the Future: Exploring the Impact of Technology and AI on Modern*

Translation Studies. New Delhi: CSMFL Publications.

- [33] Welchman, Lynn. (2000). *Beyond the Code: Muslim Family Law and the Shari'a Judiciary in the Palestinian West Bank*. Leiden: Brill.
- [34] Wu, Xiaodong, Ran Duan, & Jianbing Ni. (2024). Unveiling security, privacy, and ethical concerns of ChatGPT. *Journal of Information and Intelligence*, 2(2), 102–115.
- [35] Xu, Wei, & Xiao Tan. (2025). Blessing or cursing: Second language writing teachers' perceptions of using ChatGPT in classrooms. In Hui Cheng (Ed.), *Rethinking Writing Education in the Age of Generative AI* (pp. 71–93). Abingdon: Routledge.
- [36] Yang, Lu. (2024). Critical thinking cultivation of teachers' perspective: Conception, influencing factors and challenges in English writing teaching. *The Educational Review, USA*, 8(3), 448–453.
- [37] Zayed, Amna Ben. (2024). Evaluating the fidelity and accuracy of ChatGPT 4 and Google Translate in translating legal English documents into Arabic—and vice versa. *Faculty of Languages Journal – Tripoli, Libya*, 1(29), 63–87.

Alalddin A. Al-Tarawneh is an Associate Professor in the Department of English and Translation at Zarqa University, Jordan. He earned his Ph.D. in Translation Studies from Queen's University Belfast and joined Zarqa University's faculty in 2017, where he teaches and conducts research on translation theory and practice, with particular emphasis on issues of meaning and context. He has published multiple academic articles and book chapters and participates actively in research on language and AI in translation.

Mohammed Al-Badawi is a Full Professor of Linguistics at Alzaytoonah University of Jordan, whose expertise spans an impressive array of disciplines. With a focus on Sociolinguistics, Interlinguistics, and Anthropological Linguistics, this academic luminary brings a wealth of knowledge to the realm of language studies. Possessing a diverse skill set, including proficiency in Stylistics, Pragmatics, Applied Linguistics, Cognitive Linguistics, Discourse Analysis, Critical Discourse Analysis, Lexicography, Cognitive Semantics, Syntax, and Conceptual Metaphor Theory, this scholar is at the forefront of linguistic exploration.

Muath Algazo holds a PhD in Applied Linguistics from Western University. He is currently an assistant Professor at Al-Zaytoonah University of Jordan, where he teaches in the Department of Applied English. His research interests include language acquisition, language assessment, second language teaching and learning.

Rula Ahmad Mahmoud Abu-Elrob is an Assistant Professor in the department of English Language and Literature at Al-Zaytoonah University of Jordan. Her research interests include Conversation Analysis, Discourse Analysis, Pragmatics, Semantics and Applied Linguistics.

Bilal Alsharif holds a PhD in Linguistics from Newcastle University. He is currently an assistant Professor at Mutah University, where he teaches in the Department of English Language and Literature. His research interests include laboratory phonology, applied corpus linguistics, and machine translation.

Nader Albkower is an Assistant Professor in the Department of Translation at Mutah University, Jordan, where he teaches and engages in research on translation studies and intercultural communication. He holds a Ph.D. (with graduation from the United Kingdom) and has contributed to scholarly publications on translation, narrative and discourse in media and language contexts. Albkower's work spans topics in translation theory and practice, and he is active in academic research and publication within his field.

Enhancing English Speaking Skills of Vietnamese University Students Through ChatGPT-Aided Mind Maps

Thi Huong Thao Dang
Faculty of English, Hanoi Open University, Hanoi, Vietnam

Abstract—Speaking skills are vital because they enable learners to express their ideas, opinions, and understanding of the material presented by the teacher. Furthermore, ChatGPT shows significant benefits to support students' English-speaking learning. To find out the impact of using ChatGPT-aided mind maps on students' English-speaking performance, a classroom action research study was conducted with 30 second-year English major students at a university in Vietnam during the first semester of the 2025-2026 academic year. The data collection tools included a pre-test, a post-test, and a questionnaire consisting of 24 Likert scale items and three open-ended questions. Quantitative data were analyzed using IBM SPSS Statistics (Version 25.0), including descriptive statistics and paired sample t-tests. The result shows that (1) the ChatGPT-aided mind maps help students enhance their English performance in grammar, vocabulary, pronunciation, fluency, and discourse management. (2) Students have positive perceptions of using ChatGPT-aided mind maps for preparing and organizing ideas for speaking skills. This study demonstrated that ChatGPT-aided mind maps have beneficial impacts on students' English-speaking skills.

Index Terms—speaking skills, ChatGPT, mind maps, action research, English-speaking performance

I. INTRODUCTION

Four main factors affecting students' fluency in speaking are limited vocabulary, pronunciation, grammar, and self-confidence (Fitriani et al., 2015; Leong & Ahmadi, 2017; Jaya et al., 2022). Speaking is generally recognized as one of the most essential skills in English language learning, as it enables students to communicate effectively. However, for many learners of English as a foreign language (EFL), speaking is also one of the most difficult skills to master. In fact, they have difficulties in acquiring it successfully. Therefore, it is necessary to find out effective ways to support students in improving their speaking skills.

Buzan and Buzan (1993) emphasize that the mind map is a reflection of Radiant Thinking, and can be seen as a basic cognitive function. In addition, it is an effective graphic technique that offers a universal framework for enhancing learners' cognitive potential. It means that mind mapping is a useful technique for organizing and visualizing ideas helping learners generate, connect, and remember key concepts more easily. Mind maps provide a clear structure that supports coherence and reduces hesitation, making oral communication more effective. Mind maps can enhance memory, stimulate creativity, and improve coherence in both writing and speaking tasks.

Meanwhile, the rapid development of artificial intelligence (AI) has created new opportunities in language education. ChatGPT, an advanced language model developed by OpenAI, has been shown to support learners by generating ideas, supplying vocabulary, and offering contextualized input that enriches traditional classroom learning.

Given the benefits of ChatGPT for language learning and the limited research on ChatGPT-aided mind maps to enhance English-speaking skills, this study investigates the impact of ChatGPT-aided mind maps on students' English-speaking performance and their perceptions of using this tool for speaking preparation. In order to achieve the objectives of the study, two research questions are stated:

- (1) To what extent does the use of ChatGPT-aided mind maps improve the speaking skills of second-year English major students?
- (2) What are students' perceptions of using ChatGPT-aided mind maps for preparing and organizing ideas for speaking tasks?

II. LITERATURE REVIEW

A. Mind Maps

Dwina (2024) confirms that mind mapping is one of the instructional tools that can be used by teachers. Previous studies indicate that mind mapping is highly effective because it helps students connect ideas and become engaged in learning activities. Students benefit from being able to integrate disparate ideas, think creatively, and associate concepts (Buzan, 2018).

Mind mapping displays not only information, but also a subject's overall organization and the relative importance of its various components. Mind mapping has been utilized in language learning and teaching for a considerable time and has consistently demonstrated its advantages. Mind mapping is a flexible method that provides learners with a powerful graphic tool for organizing, recording, and gathering information. According to Buzan (2006), mind mapping is a visual tool that learners can use to generate and organize ideas, take notes, and develop concepts. Mind maps are highly suitable for helping learners generate ideas and connect new information with their prior knowledge (AlMutairi, 2015). It also makes it easier for students to brainstorm and share their thoughts. It is considered a supportive tool for students in developing their English-speaking skills. Bystrova and Larionova (2015) explain that the mind-mapping learning model employed concept maps, supported by images and keywords, to represent material in a comprehensive manner. Liu et al. (2018) argue that mind-mapping techniques are employed due to their advantages in teaching and enhancing both logical and innovative thinking. According to Buzan and Buzan (1993), the mind map has four essential characteristics: the crystallization of the subject of attention in a central image, the subject's main themes radiate from the central image as branches, the structural composition of branches with key images or keywords placed along connecting lines and the hierarchical representation of less important topics as subordinate branches, and the formation of a connected nodal structure by branches.

B. ChatGPT

Tertiary education could benefit from using ChatGPT for writing support, language learning, and research (Tran et al., 2025). ChatGPT has a multidimensional impact on education. It provides educational content, supports learning processes, and contributes to the development of essential skills (Ngo et al., 2025). Previous research suggests that integrating AI into the classroom can increase both motivation and academic performance (Afkarin & Asmara, 2024). Students who are interested in practical, experiential learning will benefit from using ChatGPT as a learning aid (Rudolph et al., 2023). ChatGPT is considered an educational and learning tool that creates both opportunities and challenges for students and instructors (Atlas, 2023). ChatGPT, in particular, has attracted considerable attention due to its ability to generate human-like language and has been suggested as a potentially valuable tool for language education (Baskara & Mukarto, 2023).

The benefits of ChatGPT are proven by the number of scholars. It has become a new trend in education, which is expected to be an effective alternative learning tool due to its advantages in providing quick, meaningful responses and saving time. Furthermore, it is considered a powerful tool to improve learning and teaching processes. AlAfnan et al. (2023) confirm that ChatGPT provides both instructors and students with many opportunities to improve the quality of their teaching and learning. For students, it is considered a potential alternative to traditional search engines, allowing them to quickly find millions of results for discussion. Using ChatGPT to support students in creating mind maps for English-speaking skills is highly beneficial for enriching vocabulary, expanding ideas, and saving time on task. However, one key consideration when applying ChatGPT to designing mind maps is how to provide effective prompts. The teacher should provide guidance on how to write prompts and utilize the information they receive from ChatGPT to enhance their mind maps.

C. Speaking Skills

Graddol (2006) emphasizes the growing importance of English as a key factor in academic achievement and career opportunities, as it is essential for global communication, mobility, and employability. Speaking as a reciprocal process of making meaning that involves receiving and processing input and producing spoken output (Brown, 1994; Burns & Joyce, 1997). According to Abd El Fattah Torky (2006), speaking is a bidirectional interactional process that involves a true communication of ideas, information, or feelings.

Ur (1996) argues that speaking is one of the four essential language skills, as people who acquire a language are referred to as language users. In fact, English-speaking competence is crucial in both learning and real-life situations, which helps students communicate successfully. In fact, many students cannot communicate in English effectively and confidently after studying it for many years. In the 21st century, being able to speak English is a key factor enabling learners to enter the world of work and become a global citizen. It is undeniable that learners who excel in English are more confident in communication and collaboration, which are vital in life. If we want students to engage in effective communication in English, we should create a supportive learning environment that encourages them to use English as frequently as possible.

Two important elements of the communication approach are fluency and accuracy. Students can improve their English communicative competence through classroom practice. Two main characteristics of speaking performance are fluency and accuracy. Fluency is the first characteristic of speaking performance, so teachers often focus on teaching it. Hedge (2000) states that fluency involves answering coherently, linking words and phrases smoothly, pronouncing sounds clearly, and applying appropriate stress and intonation. According to Hughes (2002), fluency is the learner's ability to speak intelligibly, preventing communication breakdowns that may lead listeners to lose interest. Accuracy is the second characteristic of speaking performance; therefore, learners should pay careful attention to grammatical structures, vocabulary, and pronunciation to improve their speaking accuracy. In terms of vocabulary, learners know how to select suitable words in appropriate contexts. It means that they can use words and expressions correctly. In

addition to speaking English accurately, learners should master grammatical structures, lexical resources, and phonological rules, including stress, intonation, and pitch.

Acquiring and mastering English language skills is not easy because of linguistic complexity (grammar, vocabulary, and pronunciation) and the four skills. The speaking problems were categorized into three main categories: affective-related issues, including self-confidence and anxiety; relevant difficulties with limited opportunities for English learning and speaking in class; and linguistically related problems, including fluency, grammar, vocabulary, and pronunciation (Jaya et al., 2022).

III. METHODOLOGY

A. Research Setting

This study was conducted with second-year English major students at a university in Vietnam during the first half of the 2025-2026 academic year. The English-speaking course lasts 12 weeks and uses Life Vision Intermediate plus the Student's Book as the main textbook. The English-speaking course was delivered for 2.5 contact hours on Wednesday mornings. The classroom is equipped with basic learning facilities, including whiteboards and a smart TV with an internet connection.

B. Participants

The participants in this study were 30 second-year English major students who participated in Speaking Skill Course 3 at the beginning of the academic year 2025/2026. The class consisted of 5 male and 25 female students, aged 18-19. Based on preliminary observations and the pre-test results, the majority of students demonstrated a basic level of English proficiency equivalent to Level 3 of the Vietnamese Six-Level Foreign Language Proficiency Framework, and encountered some difficulties in their English performance.

C. Research Design

The study employed the action research model of Kemmis and McTaggart (1988) to apply ChatGPT-aided mind maps to English-speaking skills for seven weeks. The model comprises four distinct stages: planning, action, observation, and reflection.

TABLE 1
THE SEVEN-WEEK IMPLEMENTATION OF CHATGPT-AIDED MIND MAPS

Week	Unit	Tasks
1	Pre-test	The teacher assessed students' English-speaking skills The teacher guided students to formulate prompts for using ChatGPT edit and refine the mind maps
2	Unit 1: From the page to the screen	Topic 1: Talk about books Topic 2: Talk about films Students created the mind maps related to the themes in class. The teacher provided feedback on the mind maps Students revised the mind maps at home based on the teacher's comments and ChatGPT support.
3	Unit 2: Back to nature	Students submitted both the original and revised versions of the mind maps for evaluation Topic 3: Talk about geographical features and landscapes Topic 4: Talk about the environment Students created the mind maps related to the themes in class. The teacher provided feedback on the mind maps Students revised the mind maps at home based on the teacher's comments and ChatGPT support.
4	Unit 3: Get well soon	Students submitted both the original and revised versions of the mind maps for evaluation Topic 5: Talk about injuries & illnesses Topic 6: Talk about healthy living Students created the mind maps related to the themes in class. The teacher provided feedback on the mind maps Students revised the mind maps at home based on the teacher's comments and ChatGPT support.
5	Unit 4: This is who we are	Students submitted both the original and revised versions of the mind maps for evaluation Topic 7: Talk about people's personalities Topic 8: Talk about family relationships Students created the mind maps related to the themes in class. The teacher provided feedback on the mind maps Students revised the mind maps at home based on the teacher's comments and ChatGPT support.
6	Revision	Students submitted both the original and revised versions of the mind maps for evaluation Students practiced English speaking with the revised mind maps.
7	Post-test	The teacher assessed students' English-Speaking Skills

D. Data Collection and Analysis

To collect valid and comprehensive information, this study employed three main instruments: a pre-test, a post-test, and a questionnaire. Each instrument was selected to serve a specific function in addressing the research objectives, ensuring both qualitative and quantitative perspectives were covered.

Pre-test and Post-test instruments: Speaking tests were administered to assess students' oral performance before and after the implementation of the mind map technique. The tests were adapted from the VSTEP (Vietnamese Standardized Test of English Proficiency) speaking test format. The students' performance was evaluated using the VSTEP speaking scoring criteria, which includes five aspects: grammar, vocabulary, pronunciation, fluency, and discourse management, are made public by the Vietnamese Ministry of Education and Training. The criteria were modified to align with the learning outcome of the Speaking 3 course. All the speaking tests were assessed by the course instructor, who was formally trained and certified as a VSTEP speaking examiner. Subsequently, the speaking test data were analyzed using SPSS 25.0.

Questionnaires were distributed to examine the impact of ChatGPT-aided mind maps on English-speaking skills (Questions 1-12) and to collect students' perceptions of ChatGPT-aided mind maps on their English-speaking performance (Questions 13-24). A Likert scale format (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree) was applied to quantify students' responses. This instrument was selected because it enabled the researcher to efficiently gather perspectives from all participants, ensuring a broad representation of students' views.

The internal consistency of the Likert scale items was assessed using Cronbach's alpha. The scale measuring the impact of ChatGPT-aided mind maps on English-speaking performance demonstrated high internal consistency (Cronbach's $\alpha = .988$, 12 items). Correspondingly, the questionnaire assessing students' perceptions of ChatGPT-aided mind maps showed excellent reliability (Cronbach's $\alpha = .984$, 12 items). The Likert-scale items were analyzed using SPSS version 25. Additionally, three open-ended questions were used to explore their views on using ChatGPT-aided mind maps to enhance their English-speaking skills. These responses were treated as qualitative data and analyzed with care, which provided deeper insights and richer qualitative data that could not be captured through tests alone.

To increase the reliability of the data obtained from the questionnaires, a pilot test was conducted with 10 participants from the cohort, and the instrument was reviewed by the English lecturer. Necessary revisions were made based on the feedback.

In summary, the use of multiple instruments not only enabled triangulation of data but also ensured a more holistic understanding of how the mind map technique contributed to improvements in students' speaking skills.

IV. RESULTS

A. Results From Pre-Test and Post-Test Scores

TABLE 2
PRE-TEST AND POST-TEST COMPARISONS OF STUDENTS' ENGLISH- SPEAKING PERFORMANCE (N = 30)

Speaking criteria	Pre-test M (SD)	Post-test M (SD)	t (29)	p	Cohen's d
Grammar	6.68 (0.70)	6.97 (0.99)	1.81	.081	.33
Vocabulary	6.83 (0.74)	7.38 (0.93)	3.78	.001	.69
Pronunciation	6.87 (0.66)	7.43 (0.92)	4.07	< .001	.74
Fluency	6.73 (0.72)	7.05 (1.03)	2.35	.026	.43
Discourse management	6.68 (0.73)	7.27 (0.96)	4.51	< .001	.82
Overall speaking	6.78 (0.66)	7.21 (0.94)	3.43	.002	.63

Note. M = mean; SD = standard deviation. Paired-samples *t*-tests were conducted. Effect sizes were calculated using Cohen's *d*. Values of $p < .05$ indicate statistical significance

As shown in Table 2, the results indicate that students' speaking performance improved after a 7-week implementation of ChatGPT-aided mind maps. In all six areas, the post-test mean scores were higher than the pre-test scores. For grammar, the mean increased slightly from 6.68 to 6.97. Vocabulary improved more clearly, from 6.83 to 7.38, showing that students used a wider range of words after the intervention. Pronunciation also improved from 6.87 to 7.43, indicating students' progress in stress, intonation, and pitch in speaking. The mean fluency score rose from 6.73 to 7.05, illustrating a minor improvement. Discourse management grew from 6.68 to 7.27, reflecting better organization and coherence. Finally, the overall speaking performance improved from 6.78 to 7.21.

The results also show that most aspects of English-speaking performance improved significantly after the intervention. Vocabulary ($t = 3.78$, $p = .001$), pronunciation ($t = 4.07$, $p < .001$), fluency ($t = 2.35$, $p = .026$), discourse management ($t = 4.51$, $p < .001$), and overall performance ($t = 3.43$, $p = .002$) all showed statistically significant gains. In contrast, grammar showed a minor, non-significant improvement ($t = 1.81$, $p = .081$). The results confirm that ChatGPT-aided mind maps had a positive impact on students' English-speaking performance.

The paired-sample *t*-test revealed significant improvements in fluency, overall speaking performance, pronunciation, vocabulary, and discourse management, with medium to large effect sizes (Cohen's $d = .43-.82$). Grammar showed a small-to-medium effect that did not reach statistical significance ($p = .081$, $d = .33$).

B. Results From Open-Ended Questions

Furthermore, thirty students were also asked to answer three open-ended questions to find out the most useful aspects, challenges, and strategies of using ChatGPT-aided mind maps to improve English-speaking skills.

What was the most useful thing about using ChatGPT to support mind map designs for speaking performance?

What challenges did you face when using ChatGPT to design mind maps?

How would you improve your speaking performance by using ChatGPT-aided Mind Maps?

For the perceived usefulness, the majority of students emphasized idea generation, while a substantial proportion highlighted vocabulary enrichment and clarity of organization. A smaller group noted the benefit of time saving. These findings suggest that students viewed ChatGPT as a helpful tool for preparing content and structuring ideas prior to speaking tasks. Students also share their perceptions of the benefits of using ChatGPT in mind-mapping. One student wrote that ChatGPT provided “more vocabulary, collocations for topics that I often find difficult to express.” (Student 2). Another emphasized, “the most useful thing was quickly generating a wide range of ideas” (Student 8), while another response highlighted “saving time” during preparation. These statements demonstrate that students primarily valued ChatGPT for idea generation, lexical enrichment, and efficiency in preparing ideas, therefore reducing the cognitive burden before speaking tasks.

In terms of challenges, the most frequently reported issue was the generic nature of ChatGPT outputs. Other concerns included unsuitable or missing vocabulary items, difficulties in understanding certain generated content, and over-dependence on the tool. A few students also mentioned the lack of personalization. One student noted that ChatGPT “used words I don’t like, some important details were missing,” (Student 9) while another admitted, “the main challenge was making the generic ideas fit my context.” (Student 15). A third explained, “try to not depend too much on ChatGPT and try to think for myself.” (Student 17) These reflections indicate that while students valued AI support, they were also aware of its limitations in personalization and the potential risks associated with over-reliance.

Regarding suggested improvement strategies, students most commonly pointed to the need to practice more and to expand vocabulary. In addition, some students mentioned the importance of clearer organization, others emphasized the need to build confidence, and a few highlighted the use of linking words to enhance fluency and cohesion. One student explained, “How to use linking words,” while another emphasized, “clearer organization – the mind map helps me to structure ideas better.” (Student 24). Others expressed the intention to “practice more” or “use more vocabulary and linking words.” (Student 28). These comments suggest that students perceive ChatGPT-aided mind maps as a scaffold for better organization and lexical richness; however, they still view active practice and self-regulation as essential to improving their speaking performance.

C. Results From 24 Likert Scale Items

The findings in Table 3 indicate that using ChatGPT-aided mind maps is a highly effective way for students to improve their English-speaking skills. Vocabulary-related items gain the highest mean scores with expressing ideas more precisely ($M = 4.07$) and using a wider range of words ($M = 4.03$). Students also had improvements in coherence ($M = 3.93$), grammar use ($M = 3.93$), and reduced hesitation ($M = 3.93$). Lower mean scores than previous tests were observed in pronunciation ($M = 3.67$), naturalness ($M = 3.63$), and overall performance ($M = 3.53$), yet the students' improvement was still evident.

TABLE 3
THE IMPACTS OF CHATGPT-AIDED MIND MAPS ON ENGLISH-SPEAKING SKILLS (N = 30)

Statements	Mean	SD
1. Using ChatGPT-aided mind maps helped me speak more fluently in English.	3.80	.76
2. The mind maps improved the coherence and logical flow of my ideas during speaking.	3.93	.87
3. I hesitated less when speaking because the mind maps gave me clear structure.	3.93	.87
4. ChatGPT provided me with useful vocabulary and collocations for speaking tasks.	3.87	.86
5. I could use a wider range of words in my speaking thanks to ChatGPT suggestions.	4.03	.72
6. The vocabulary from ChatGPT-aided mind maps helped me express my ideas more precisely.	4.07	.79
7. ChatGPT-supported mind maps helped me use more varied grammatical structures.	3.93	.79
8. I made fewer grammatical mistakes when I prepared with ChatGPT-aided mind maps.	3.77	.82
9. With clearer preparation from mind maps, I felt more confident in my pronunciation.	3.67	.71
10. I spoke more naturally and clearly after preparing with ChatGPT-aided mind maps.	3.63	.77
11. ChatGPT-aided mind maps improved my overall performance in speaking tasks.	3.53	.78
12. I believe this strategy will help me perform better in real speaking tests.	3.77	.77

Table 4 presents learners' perceptions of using ChatGPT-supported mind maps. It highlights that students expressed positive perceptions of ChatGPT-aided mind maps. The highest level of agreement was found in perceived usefulness in preparing for speaking tasks ($M = 4.20$) and increased confidence in speaking ($M = 4.00$). Students also valued the tool for providing new vocabulary ($M = 3.87$) and supporting the organization of ideas ($M = 3.83$). Moderate means were reported for motivation to practice ($M = 3.67$) and concerns about over-reliance on ChatGPT ($M = 3.80$). Overall, students indicated willingness to continue using and recommending the strategy ($M \approx 3.80$ – 3.83).

TABLE 4
STUDENTS' PERCEPTIONS OF USING CHATGPT-AIDED MIND MAPS FOR ENGLISH-SPEAKING PREPARATION (N = 30)

Statements	Mean	SD
1. ChatGPT-aided mind maps were useful in preparing for my speaking tasks.	4.20	.61
2. The use of ChatGPT saved me time when generating ideas for speaking.	3.77	.57
3. ChatGPT-aided mind maps helped me organize my ideas more effectively.	3.83	.65
4. ChatGPT provided me with new vocabulary and expressions that I could use in speaking.	3.87	.57
5. The use of ChatGPT increased the variety of ideas I could include in my speech.	3.73	.64
6. ChatGPT-aided mind maps helped me feel more prepared for speaking tasks.	3.83	.65
7. Using ChatGPT-aided mind maps motivated me to practice speaking more regularly.	3.67	.71
8. I felt more confident when speaking after preparing with ChatGPT-aided mind maps.	4.00	.70
9. I sometimes relied too much on ChatGPT instead of using my own ideas.	3.80	.71
10. ChatGPT's suggestions were not always fully suitable, and I needed to adapt them.	3.77	.68
11. I would like to continue using ChatGPT-aided mind maps for future speaking practice.	3.83	.65
12. I would recommend ChatGPT to other students to support their speaking preparation.	3.80	.76

V. DISCUSSION

The findings indicate that ChatGPT-aided mind maps have a positive impact on students' English-speaking skills. Notable improvements were recorded in three aspects of English speaking: discourse management, vocabulary, and pronunciation, while fluency progressed more slowly, and grammar showed the smallest improvement. This can be explained by the fact that ChatGPT provided students with useful vocabulary, linking words, and sentence patterns. With ChatGPT support, it is easy for students to find useful key points and lexical resources, helping them remember and apply them more effectively during speaking tasks.

The results of this action research provide crucial evidence that integrating ChatGPT-aided mind maps had a positive impact on the speaking performance of English-major students. Quantitative findings from the pre- and post-tests demonstrated great improvements in discourse management, vocabulary, and pronunciation, with less dramatic but still valuable gains in fluency and grammar. This can be attributed to the cognitive scaffolding provided by ChatGPT. Based on the prompts provided by students, ChatGPT offered a range of lexical items with their functions in authentic contexts, allowing them to choose appropriate keywords to create their own mind maps. Vocabulary expansion helped them to generate lexical sets, collocations, and synonyms that are very useful for English-speaking preparation. Similarly, the pronunciation improvement may be explained by the fact that ChatGPT provided vocabulary with meaning explanations and model sentences, which were crucial for applying effectively in their speaking practice and contributed to clearer articulation and intonation. The great improvement in discourse management also reflects the organizational benefits of mind mapping, which allowed learners to plan ideas systematically and present them coherently. In fact, ChatGPT suggested a wide range of linking words, along with detailed explanations and examples, to convey clear, logical ideas. However, the minor progress in grammar and fluency indicates that while ChatGPT provides useful input for preparation, it does not replace the need for spontaneous interaction, which is essential for mastering real-time accuracy and smoothness of delivery. Thus, while the quantitative data confirm that ChatGPT-aided preparation has noticeable improvements on linguistic performance, they also highlight the continuing importance of communicative practice in reinforcing these gains. The results from the questionnaires also showed that ChatGPT-aided mind maps assisted students in enriching their vocabulary the most, particularly with collocations and a wider range of words, enabling them to express their ideas more precisely. Additionally, the pronunciation was improved through their natural speaking. Discourse management was also enhanced with logical flow and coherence in speaking. Students also found that ChatGPT provided them with more varied grammar structures, resulting in fewer mistakes during speaking time. However, they need to practice more to perfect it. They confirmed the improvement in fluency due to the support of ChatGPT in creating mind maps.

The students have positive perceptions of using ChatGPT-aided mind maps in English speaking skills. As a result, they are confident in speaking English, with a richer vocabulary, clearer pronunciation, and well-organized ideas. The results suggest that ChatGPT-aided mind maps enhanced students' preparation, confidence, and lexical resources for speaking tasks. The tool was particularly effective in boosting confidence, likely due to structured support and accessible input. However, students also noted risks of over-reliance and the need to adapt ChatGPT suggestions. These findings highlight both the benefits and limitations of ChatGPT-assisted strategies, underscoring the importance of striking a balance between ChatGPT support and independent idea generation to optimize learning outcomes.

Questionnaire responses confirmed that students show positive perceptions of using ChatGPT-aided mind maps in English-speaking skills. The highest level of agreement is usefulness in speaking preparation and increasing confidence, which is explained by the fact that when students have enough vocabulary, linking words, and grammar structures to express their ideas, they are willing to speak. The more they speak, the better their pronunciation and fluency are. The qualitative data add depth to these interpretations by revealing how students perceived the benefits and challenges of ChatGPT-aided mind maps. Furthermore, the results showed overwhelmingly positive perceptions: all participants reported that they enjoyed learning with ChatGPT, and the majority indicated that the tool made them more motivated and confident. This suggests that the affective benefits of ChatGPT support played an important role in enabling students to improve their performance. When learners are less anxious and more willing to take risks, they are more

likely to apply new vocabulary and experiment with pronunciation, which in turn facilitates faster skill improvement. Thanks to the benefits of ChatGPT support, students are motivated to practice English and are willing to use it in future learning. However, they also mentioned one concern about over-reliance on ChatGPT and the generic nature of ChatGPT outputs, which can reduce their critical thinking and make them confused. Although ChatGPT can enrich language learning, it cannot substitute for students' own critical thinking, creativity, and communicative practice. ChatGPT is a powerful facilitator; however, it is essential for students to recognize the drawbacks of using ChatGPT to understand how to use it properly and effectively in their learning progress with the guidance of teachers. Taken together, the combination of quantitative and qualitative findings suggests that ChatGPT-aided mind maps are most effective as a scaffold for better organization and lexical richness; however, students still need active practice and self-regulation as essential to improving their speaking performance.

Both ChatGPT and mind maps are effective to support English speaking performance. These findings align with Buzan and Buzan's (1993) claim that visual mapping enhances the logical structuring of ideas. They also claimed that clearer thinking and information memory were promoted by mind mapping, while Bystrova and Larionova (2015) and Liu et al. (2018) argued that digital mind maps improve the logical structuring of ideas and comprehensive materials. The present study confirms earlier findings showing that ChatGPT-aided mind maps enabled students to organize coherent ideas in their speaking performance, as reflected in significant improvements in discourse management scores.

The study also matches with prior research on the role of technology and CALL tools in language learning. Ur (1996) emphasized the importance of providing learners with confidence and motivation in speaking tasks, while Burns and Joyce (1997) stressed that both cognitive and affective support should be integrated in speaking instruction. The questionnaire results affirmed that both of these dimensions were supported by ChatGPT. For the cognitive aspect, ChatGPT supplied lexical resources and sentence structures. From an affective perspective, it reduced students' anxiety and motivated them in English speaking performance.

In terms of specific speaking challenges, Jaya et al. (2022) found that even English majors in Indonesia encountered difficulties in fluency, grammar, vocabulary, and pronunciation during speaking performance. Similarly, Leong and Ahmadi (2017) reported vocabulary limitations, lack of confidence, and poor discourse management as recurring challenges for EFL learners. The current research confirms these findings in the Vietnamese context but takes it further by showing that ChatGPT-aided mind maps can moderate these challenges. It is clear that the improvements in students' lexical items, discourse management, pronunciation, grammar, and fluency directly address the problems, which demonstrates that ChatGPT aided-mind maps reduce affective barriers highlighted in previous studies.

Baskara and Mukarto (2023) investigated ChatGPT's role in higher education and confirmed that using it appropriately can improve vocabulary acquisition, stimulate idea generation, and foster critical thinking. The current study elaborates on their work by focusing specifically on English-speaking performance and by integrating ChatGPT into mind map design rather than using it as a separate text generation tool. This result highlights the importance of combining ChatGPT with well-established teaching strategies to achieve meaningful learning outcomes. Moreover, while Baskara and Mukarto (2023) expressed concerns about the generic nature of ChatGPT outputs, the present study found similar challenges, as some students reported difficulties in adapting ChatGPT's support into English performance. This alignment strengthens the need for guided, critical use of ChatGPT in English-speaking skills.

At the same time, the minor improvement in grammar and fluency observed in this study aligns with previous CALL research that emphasizes the limits of technological support for spontaneous oral production. Although ChatGPT is very effective at improving English speaking skills, it requires ongoing practice and feedback to facilitate sustained improvement in fluency among EFL learners. As Brown (1994) and Hughes (2002) argued, communicative competence develops through authentic interaction, negotiation of meaning, and real-time processing. Although ChatGPT provided valuable input during the preparation phase, it could not fully reproduce the dynamics of authentic communication.

In summary, the present study supports earlier findings on the benefits of mind mapping, CALL tools, and affective support in speaking instruction while adding new evidence on the potential of AI integration in EFL contexts. It confirms that ChatGPT-aided mind maps can enhance discourse management, vocabulary, and pronunciation, which echoes established theories while extending them into the emerging field of AI-assisted pedagogy. At the same time, it highlights limitations identified in previous research, particularly concerning generic output and over-reliance on technology. By bridging traditional strategies with innovative AI tools, this study contributes to the evolving conversation on how technology can best support learners' speaking development in higher education.

VI. CONCLUSION

This study demonstrated that the use of ChatGPT-aided mind maps helps students improve their English-speaking skills. The pre-test and post-test results indicated marked progress in vocabulary, pronunciation, fluency, discourse management, and overall performance, although improvements in grammar were only slight. Questionnaire data further confirmed that students agreed with the usefulness of the tool for preparing, organizing ideas, and building confidence. These findings from open-ended responses shared the same points as those of the two previous ones. Students emphasized the value of ChatGPT in generating a wide range of ideas, enriching vocabulary, and saving time. However, qualitative insights highlighted concerns about generalized outcomes, a lack of personalization, and the risk of over-

reliance, suggesting that while ChatGPT is a powerful support tool, its effectiveness depends on its use with guidance and critical awareness.

The findings provide insights for both teaching and learning. For teachers, integrating ChatGPT-aided mind maps can be an effective classroom strategy to provide students with structured support, richer vocabulary, and clearer idea organization before speaking tasks. For students, the tool can serve as a means of self-study, encouraging autonomy, practice, and confidence. However, to avoid over-reliance on ChatGPT, it should be combined with communicative activities such as pair work, group discussions, and presentations, and teachers' clear guidance on writing prompts. A balanced approach can therefore maximize the benefits of ChatGPT support while still fostering creativity, spontaneity, and independent language use.

Although the study's results demonstrate the positive impact of using ChatGPT-aided mind maps to enhance students' English-speaking skills, it has certain limitations in scope and generalizability because it was conducted with a single class of 30 students over seven weeks. Further studies should extend the intervention time and utilize a large sample size to provide more comprehensive contributions of ChatGPT-aided mind maps to students' improved English skills.

REFERENCES

- [1] Abd El Fattah Torky, S. (2006). *The effectiveness of a task-based instruction program in developing the English language speaking skills of secondary stage students* [Doctoral dissertation]. Ain Shams University.
- [2] Afkarin, M. Y., & Asmara, C. H. (2024). Investigating the Implementation of ChatGPT in English Language Education: Effects on Student Motivation and Performance Levels. *Journey: Journal of English Language and Pedagogy*, 7(1), 57-66. <https://doi.org/10.33503/journey.v7i1.865>
- [3] AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: opportunities, challenges, and recommendations for communication, business writing, and composition courses, *Journal of Artificial Intelligence and Technology*, 3(2), 60-68. <https://doi.org/10.37965/jait.2023.0184>
- [4] AlMutairi, A. N. (2015). The effect of using brainstorming strategy in developing creative problem-solving skills among male students in Kuwait: A field study on Saud Al-Kharji School in Kuwait City. *Journal of Education and Practice*, 6(3), 136-145.
- [5] Atlas, S. (2023). *ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI*. College of Business Faculty Publications. https://digitalcommons.uri.edu/cba_facpubs/548
- [6] Baskara, F. R., & Mukarto, F. X. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 7(2), 343-358. <https://doi.org/10.21093/ijeltal.v7i2.1387>
- [7] Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. Prentice Hall Regents.
- [8] Burns, A., & Joyce, H. (1997). *Focus on speaking*. Sydney, Australia: National Center for English Language Teaching and Research.
- [9] Buzan, T. (2006). *The mind map book*. London, England: Pearson Education.
- [10] Buzan, T. (2018). *Mind map mastery: The complete guide to learning and using the most powerful thinking tool in the universe*. Watkins Media Limited.
- [11] Buzan, T., & Buzan, B. (1993). *The mind map book*. Dutton.
- [12] Bystrova, T., & Larionova, V. (2015). Use of virtual mind mapping to effectively organise the project activities of students at the university. *Procedia – Social and Behavioral Sciences*, 214, 465-472. <https://doi.org/10.1016/j.sbspro.2015.11.724>
- [13] Dwina, L. (2024). ChatGPT: Empowering self-directed writing through mind mapping and AI-assisted composition. In *The 14th Annual International Symposium of Foreign Language Learning* (pp. 33-45). NST Proceedings. <https://doi.org/10.11594/nstp.2024.3804>
- [14] Fitriani, D. A., Apriliaswati, R., & Wardah. (2015). A study on students' English-speaking problems in speaking performance. *Jurnal Pendidikan dan Pembelajaran Untan*, 4(9), 1-13.
- [15] Graddol, D. (2006). *English Next: Why Global English May Mean the End of 'English as a Foreign Language'*. British Council.
- [16] Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- [17] Hughes, R. (2002). *Teaching and researching speaking*. New York, NY: Pearson Education.
- [18] Jaya, H. P., Petrus, I., & Pitaloka, N. L. (2022). Speaking performance and problems faced by English major students at a university in South Sumatera. *Indonesian EFL Journal*, 8(1), 105-112. <https://doi.org/10.25134/ieflj.v8i1.5603>
- [19] Kemmis, S., & McTaggart, R. (1988). *The action research planner* (3rd ed.). Geelong: Deakin University Press.
- [20] Leong, L.-M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English-speaking skill. *International Journal of Research in English Education*, 2(1), 34-41. <https://doi.org/10.18869/acadpub.ijree.2.1.34>
- [21] Liu, Y., Tong, Y., & Yang, Y. (2018). The application of mind mapping into college computer programming teaching. *Procedia Computer Science*, 129, 66-70. <https://doi.org/10.1016/j.procs.2018.03.047>
- [22] Ngo, Cong-Lem., Ali, S., & Diki, T. (2025). A Systematic Review of the Limitations and Associated Opportunities of ChatGPT. *International Journal of Human-Computer Interaction*, 41(7), 3851-3866. <https://doi.org/10.1080/10447318.2024.2344142>
- [23] Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- [24] Tran, T. T. P., Dang, T. N., & Nguyen, V. L. P. (2025). Master students' perceptions of how ChatGPT influenced critical thinking in academic writing at the Industrial University of Ho Chi Minh City. *International Journal of AI in Language Education*, 2(2), 20-39. <https://doi.org/10.54855/ijaile.25222>
- [25] Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge, England: Cambridge University Press.

Thi Huong Thao Dang is a lecturer at the Faculty of English, Hanoi Open University, Vietnam. She has been teaching English and supervising undergraduate and postgraduate students for 23 years. She has more than 15 years of experience as a master trainer for teachers at primary, secondary, and high schools across different areas of Vietnam. She completed an MA in English Teaching Methodology at the University of Languages and International Studies (ULIS). She received her PhD in contrastive linguistics in 2018. Furthermore, she has presented at numerous national and international TESOL conferences, including Asia TEFL, Engaging VietNam, Thai TESOL, CamTESOL, and VietTESOL. Her research interests are in ELT methodology, learner autonomy, ICT, and contrastive linguistics. She can be contacted via email at thaomanhutehy@gmail.com

EFL Policy Alignment With Saudi Arabia's Economic Transformation and Vision 2030: A Systematic Review

Saad Eid Albaqami

Department of Foreign Languages, Faculty of Arts and Humanities, Al-Baha University, Saudi Arabia

Dhaif Alzahrani

King Khalid University, Abha, Saudi Arabia

Mohammad Jamshed*

English Department, College of Science & Humanities, Prince Sattam Bin Abdulaziz University, Saudi Arabia

Mohammad Rezaul Karim

English Department, College of Science & Humanities, Prince Sattam Bin Abdulaziz University, Saudi Arabia

Abstract—Saudi Arabia's Vision 2030 aims at economic diversification and the establishment of a knowledge-based economy, emphasizing the importance of EFL education for global integration and engagement. This systematic review examined 44 studies on EFL policy alignment with Vision 2030's economic objectives, using the PRISMA framework and data from various academic databases, covering the period from 2013 to 2026. The findings revealed a substantial alignment with Vision 2030, highlighting the expansion of EFL programs, integration of technology, and the strategic role of English proficiency in enhancing global competitiveness. However, the study noted significant challenges to achieving educational objectives, including sociocultural barriers, insufficient teacher preparedness, outdated assessments, inadequate feedback, insufficient infrastructure, and curriculum misalignment, particularly regarding the underrepresentation of females in prestigious roles. The study advocates for a focus on vocational linguistic competence in EFL education to support economic transformation, recommending tailored professional development for teachers and improved integration of soft skills into the curriculum.

Index Terms—curriculum misalignment, global competitiveness, human capital, knowledge-based economy, skilled workforce

I. INTRODUCTION

Mastering English is essential for professionals across sectors such as business, science, technology, and diplomacy, as it enables global communication and collaboration. In international business, proficiency in English improves interactions, increases job opportunities, enhances operational efficiency, and offers a competitive advantage (Agustina et al., 2024; Junyi, 2024; Barnawi & Al-Hawsawi, 2016; Abro et al., 2025; Alneyadi et al., 2023). English is a crucial global language essential for various interactions. The rising need for skilled speakers highlights the importance of quality education for English language learners worldwide. Saudi Arabia's Vision 2030 Human Capability Development Program (2022) emphasizes English proficiency to enhance foreign investment and drive economic growth, particularly in tourism and investment sectors, in line with UN and OECD goals for higher education and innovation (OECD, 2019; Abbott & Lee, 2022). Significant educational reforms include the introduction of English in primary schools from Grade 1, the implementation of mandatory English tracks in technical colleges, comprehensive teacher retraining programs, and alignment with the Common European Framework of Reference for Languages (CEFR). Effective EFL policies are essential for realizing Vision 2030's economic transformation goals, as they improve English proficiency, which enhances human capital, attracts foreign investment, boosts global trade, ensures economic sustainability, and fosters global engagement (Ndiaye et al., 2024; Al-Mwzaiji & Muhammad, 2023; Althobaiti, 2025; Ismail & Namaziandost, 2025; Strango et al., 2025).

Multiple studies have highlighted the need to align economic goals with educational initiatives through a competency-based approach that integrates language and economic education. Language and economic literacy are essential for fostering economic change, as highlighted in the State of Learning and Skills 2030/2040 study. A study by Mohiuddin et al. (2023) highlighted the need for reforms in higher education, promoting modern curricula, industry-relevant outcomes, skilled graduates, faculty development, and international collaborations to meet market demands and

* Corresponding Author.

foster a knowledge-based society. Alfahadi (2019) examined the influence of Saudi Vision 2030 on EFL teaching in Saudi public universities, finding a positive impact on educational development and highlighting challenges in adopting advanced teaching methods based on a survey of 500 EFL teachers. Samarin and Al-Asfour (2023) emphasized the role of socio-political factors in human resource development, highlighting workforce adaptability and early career exploration. Dhawi and Albaqami (2017) called for enhanced university infrastructure to meet economic demands. Allmnakrah and Evers (2020) advocated for educational reforms aligned with Saudi Vision 2030, noting that existing policies limit language acquisition, impacting English proficiency and economic diversity.

However, some studies highlighted misalignment and inconsistencies between EFL policies and human capability strategies. For example, Banafi (2025) highlighted the need for better alignment between university language policy and human development, as unclear outcomes from recent educational reforms in Saudi Arabia may hinder effective policy implementation and the development of essential skills for economic and social progress. Asem et al. (2024) analyzed digital transformation in Saudi organizations under Vision 2030, focusing on strategies, innovations, and challenges related to economic diversification and social development, while highlighting a notable gap between policy directives and improvements in human capabilities. A study by Karim and Hasan (2024) found that many EFL instructors in Saudi Arabia were undertrained and unmotivated to effectively use technology, preferring traditional teaching methods. Furthermore, Abdullateef et al. (2023) highlighted misalignment between educational outcomes and labor-market demands across five Saudi universities, advocating curriculum reforms to better equip students with skills relevant to six significant gig projects. Singh et al. (2022) pointed to the critical role of education and training in fostering economic growth in line with the UN Sustainable Development Goals and Saudi Vision 2030. A mixed-method study by Saleh Alhamlan (2025) involving 760 students found that supportive professors and positive learning environments enhanced English proficiency, while issues such as large class sizes and language anxiety posed challenges.

Some studies indicate that EFL assessment policies are misaligned with the goal of improving communicative and social skills as outlined in Vision 2030. Research by Latif and Wasim (2024) found that Saudi tertiary EFL instructors favored standard assessment methods over constructivist approaches, highlighting the necessity for legislative reform, the creation of classroom-based assessment frameworks, and enhanced assessment literacy through professional development. Similarly, Latif (2021) explored the assessment practices of tertiary EFL educators in Saudi Arabia, emphasizing a tendency towards constructivist methods influenced by sociocultural theory and pragmatism, while also noting the persistence of traditional assessment techniques and significant gaps in teachers' assessment literacy. Rauf and McCallum (2020) highlighted deficiencies in assessment literacy among EFL teachers that hindered proper alignment of assessments with student learning, and suggested professional development to enhance teachers' assessment skills in line with Vision 2030. Additionally, Aburizaizah (2022) emphasized that centralized bureaucracy adversely affected creativity and report reliability in four Saudi public universities, noting that although accreditation set quality standards, it failed to cover the entire quality assurance framework in the country. Alharbi (2024) emphasized that traditional assessment methods hindered the implementation of Communicative Language Teaching in Saudi EFL classrooms, highlighting the need for improved English assessments, curriculum updates, teacher training, and test adaptations that promote communicative learning.

The current framework of continuing professional development (CPD) in Saudi Arabia is inadequately aligned with the innovation goals outlined in Saudi Vision 2030. A study by Abdullateef et al. (2023) identified significant discrepancies between existing educational curricula and teaching techniques in Saudi universities, indicating that CPD initiatives were not equipping professionals with the necessary skills to support the country's transformative objectives. This misalignment is particularly critical in the energy sector, where there is an urgent need for skill development aligned with Vision 2030. Research by Alhamad et al. (2024) advocated integrating technological competencies into CPD programs to enhance sustainable energy development strategies. Furthermore, Alnaser et al. (2024) conducted a SWOT analysis across 20 universities, revealing the need to reform architectural education to include more sustainability and digital application courses, highlighting a shortfall in Industry 4.0 training. Innovative training methods, such as virtual reality (VR), are being recognized for their potential to enhance leadership and communication skills, particularly among women leaders (Fryan et al., 2025). Additionally, the growth of the digital economy, fueled by the adoption of application programming interfaces (APIs), stressed the necessity for ongoing skill development, as noted by Ali et al. (2025). Aligning API-related strategies with national objectives is essential to support Saudi Arabia's digital advancement, further asserting the pivotal role of CPD initiatives in cultivating digital literacy and technical expertise in a rapidly changing digital environment.

II. METHODOLOGY

A. Research Questions

- A. How do Saudi EFL policies align with Vision 2030 economic objectives in the literature?
- B. How does English language planning support human capital development and workforce readiness?
- C. How do pedagogical reforms in EFL contribute to national economic competitiveness?
- D. How is Saudi EFL policy ideologically framed within economic reform?

B. Research Design

The research conducted a systematic review examining the alignment of EFL policies and language-planning initiatives in Saudi Arabia to support economic transformation under Vision 2030. Utilizing the PRISMA 2020 framework, the study defined its purpose, developed a protocol, conducted literature searches, systematically screened data, collected information, and synthesized results through thematic analysis.

C. Search Strategy

An extensive literature search was conducted through major academic databases, including Scopus, Web of Science, ERIC, ScienceDirect, SpringerLink, Taylor & Francis, and Wiley Online Library, as well as Google Scholar. This systematic approach aimed to gather a wide range of studies on English as a Foreign Language (EFL) policy and language education in Saudi Arabia, with a specific focus on Vision 2030. A structured search strategy was utilized, incorporating keywords, descriptors, and Boolean operators, and targeting core terms such as "EFL policy," "English language education," and "language planning." The search included studies published in English, with no restrictions on publication year.

D. Eligibility Criteria

Inclusion and exclusion criteria were defined to guarantee the relevance and quality of the studies reviewed. Studies were included if they:

- (a) Focused on EFL policy and education, and language-planning initiatives in Saudi Arabia;
- (b) Addressed economic development, workforce readiness, globalisation, vision 2030 objectives;
- (c) Accepted publication formats include peer-reviewed journal articles, scholarly book chapters, and doctoral dissertations;
- (d) Presented empirical findings, policy analysis, and theoretically grounded discussion.

Studies were excluded if they:

- (a) Focused exclusively on contexts outside Saudi Arabia;
- (b) Addressed English language teaching practices without policy-level or socio-economic relevance;
- (c) Were non-academic publications such as editorials, opinion pieces, reports, or conference abstracts; or
- (d) Lacked sufficient methodological or analytical clarity.

E. Screening Process

The screening process used a structured, multi-stage approach to systematic study selection. A total of 512 records were first identified. After eliminating 168 duplicates, 344 unique studies were left. Screening of titles and abstracts resulted in the exclusion of 252 records, leaving 92 full-text articles for evaluation. Of these, 48 studies were excluded for insufficient policy focus or relevance, leaving 44 included in the systematic review. This process was conducted in accordance with PRISMA guidelines to maintain rigor and transparency.

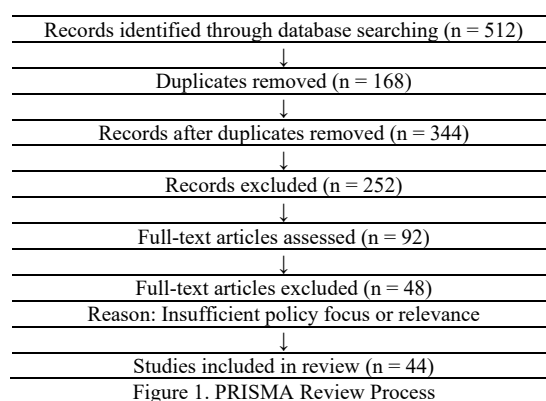


TABLE 1
ARTICLES INCLUDED IN THE SYSTEMATIC REVIEW (N = 44)

S. No.	Author(s)	Year	S. No.	Author(s)	year
1	Faruk	2013	23	Alqahtani, M. H.	2022
2	Alshammari, A. K.	2015	24	Ismail & Kassem	2022
3	Al-Zahrani & Rajab	2017	25	Al-Seghayer, K.	2023
4	Yusuf	2017	26	Almohammadi, M.	2023
5	Elyas & Badawood	2017	27	Alghamdi et al.	2023
6	Alzahrani, A. K.	2017	28	Almalki et al.	2023
7	Mitchell & Alfuraih	2017	29	Samarin & Al-Asfour	2023
8	Alqahtani, S. M. A.	2018	30	Ahmed, M. R. A	2024
9	Alanazi & Widin	2018	31	Alainati and Al-Hunaiyyan	2024
10	Sharma, V.	2018	32	Khreisat et al.	2024
11	Alnasser	2018	33	Almesaar, O.	2024
12	Al-Ahdal & Al-Awaid	2018	34	Elyas & Al-Hoorie	2024
13	Al Mukhallafi, T. R.	2019	35	Afzal, N.	2024
14	Alamr, A. S.	2019	36	Al Fraidan & Alaliwi	2024
15	Almoaily, M.	2019	37	Alqahtani, M. H.	2024
16	Al-Tamimi, R.	2019	38	Althobaiti, A.	2025
17	Alfahadi, A. M.	2019	39	Aloairdhi et al.	2025
18	Al-Shehri, S.	2020	40	Al Shammre, A. S.	2025
19	Alzhrani & Alkubaidi	2020	41	Alenazi, Y.	2025
20	Alfuhaid, S. R.	2021	42	Alkhalil, S. F. A.	2025
21	Albiladi, W. S.	2022	43	Alzahrani, A. M.	2026
22	Khawaji, A.	2022	44	Chowdhury et al.	2026

F. Data Extraction and Coding

A structured extraction framework was employed to systematically review and extract data on authors, publication years, research focuses, methodological approaches, policy contexts, and contributions to English as a Foreign Language (EFL) policy, language planning, and economic transformation. The extracted data underwent thematic coding, utilizing a deductive-inductive analysis method. Initial codes were derived from the data and the conceptual focus on Vision 2030, leading to the formulation of higher-order codes. This cyclic coding method enabled the identification of prevalent themes and convergence and divergence within the analyzed corpus.

G. Data Synthesis

The synthesis of extracted data employed narrative thematic synthesis for systematic reviews across diverse study designs, focusing on interpretation rather than statistical aggregation. Findings were organized into four key themes: strategic policy alignment, human capital development, institutional and pedagogical implementation, and sociocultural and ideological dimensions of EFL policy under Vision 2030.

III. RESULT AND DISCUSSION

In this systematic review, 44 studies were examined to assess how EFL policies and language-planning initiatives in Saudi Arabia aligned with the economic transformation objectives of Vision 2030. The review used narrative thematic synthesis to identify overarching patterns and distinctions across several areas, including policy analysis, institutional practices, pedagogical reforms, and sociocultural factors. The results highlight four key thematic areas that encompass the strategic, educational, and sociocultural aspects of EFL policy in relation to national economic priorities. For a detailed illustration of the year-wise publication trend, see Figure 2.

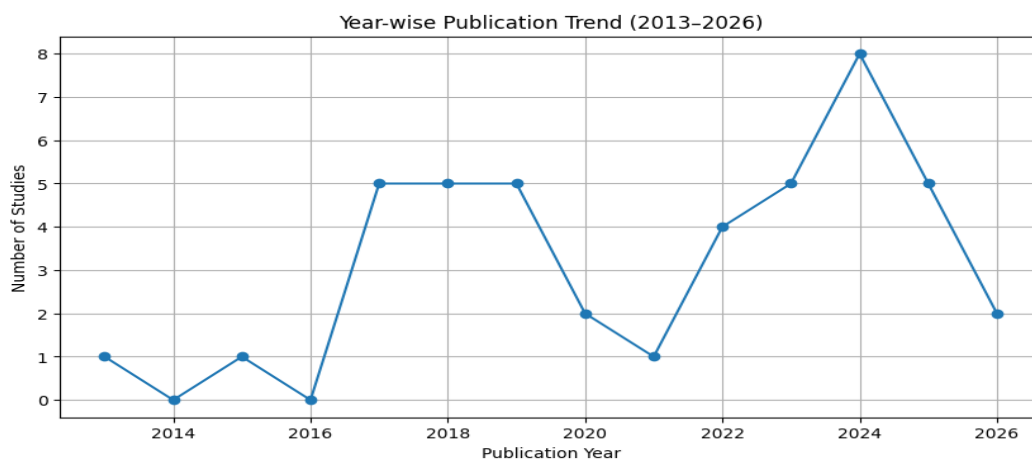


Figure 2. Year-Wise Publication Trend of Studies on EFL Policy Alignment With Saudi Vision 2030 (2013–2026)

A. Repositioning EFL Policies for Vision 2030–Driven Economic Development

Aligning EFL policies with Saudi Arabia's Vision 2030 is crucial to economic transformation. Educational reforms and language policies should enhance human capital for diversification and technological progress while reflecting cultural contexts and national priorities. Reviewed studies suggest that EFL policy in Saudi Arabia now aligns with the country's economic goals under Vision 2030, framing English as a tool for economic diversification and national development rather than just an academic discipline (Alzhrani & Alkubaidi, 2020; Alfahadi, 2019; Omer et al., 2024; Banafi, 2025; Khawaji, 2022). The studies highlight a connection between English proficiency and enhanced employability and productivity, supporting Saudi Arabia's Vision 2030 economic modernization objectives. English is recognized as a tool for empowerment and global mobility, particularly among youth and urban populations (Jabeen, 2023; Althobaiti, 2025). Bahanshal (2023) explored the shift from teaching English as a Foreign Language to English as a Second Language in Saudi Arabia, aligned with Vision 2030. Findings suggested that modernizing English language education strategies in higher education might enhance learners' skills and proficiency. Ahmed (2024) noted a shift from a symbolic language policy to an instrumental approach, in which English is seen as a key driver of global integration. Economic factors primarily drive the growth of English-medium instruction in Saudi institutions of learning, including international accreditation, improved research competitiveness, and better alignment with labor-market needs (Elyas & Al-Hoorie, 2024; Alfahadi, 2019; Almohammadi, 2023). Stakeholders acknowledge both the economic rationale and the strategic importance of English as a Foreign Language (EFL) reforms; however, disparities in resources and support lead to inconsistent policy implementation across educational contexts. Some studies (Alhazmy, 2025; Al Mukhallafi, 2019; Ndiaye et al., 2024; Alfahadi, 2019) have identified issues such as unclear curriculum objectives, challenges in self-assessment and guidance, sociocultural and implementation barriers, and difficulties in adopting high-tech teaching methods that hinder progression. Overall, the findings indicate robust alignment between EFL policy and the economic goals outlined in Vision 2030, yet significant operational challenges continue to hinder its practical realization. This discussion answers the first research question.

B. Language Planning and English for Human Capital and Workforce Readiness

Saudi language planning emphasizes English as essential for developing human capital and workforce readiness, aligning with Vision 2030 and the Human Capability Development Program. English proficiency improves skill development and job competitiveness, supported by an educational framework that fosters a skilled, adaptable, and socially responsible workforce (Abubakar et al., 2024; Banafi, 2025; Fraidan & Alaliwi, 2024; Alzahrani, 2026; Ismail & Namaziandost, 2025). Saudi language planning initiatives have enhanced EFL curricula by emphasizing communicative competence, professional discourse, personalized learning, culturally responsive pedagogy, and transferable skills, aiming to better equip learners for workforce readiness and economic participation (Jabeen, 2023; Banafi, 2025; Ndiaye et al., 2024; Al-Awaid & Hussain, 2025; Fallatah, 2021). Recent curriculum reforms, as noted by Albiladi (2022) and Khawaji (2022), have emphasized critical thinking, collaboration, and effective communication, aligning language education with labor-market expectations. Yusuf (2017) and Almoaily (2019) highlighted that the reforms in the English curriculum at Saudi universities aimed to prepare graduates for employment in the private sector and multinational environments, with both instructors and students acknowledging the role of English in improving career opportunities, especially in science, technology, and business. Students show greater motivation and find English learning more relevant when it is linked to employability, professional mobility, and future economic prospects (Zawawi et al., 2025; Yang, 2017; Alnatheer, 2013; Sharma, 2018; Alzahrani, 2017; Alnasser, 2018). These findings highlight that the language planning frameworks of Vision 2030 focus on workforce readiness, emphasizing the importance of learner experiences and institutional conditions. English is critical for improving workforce quality, but its success depends on ongoing pedagogical support, learner engagement, and institutional consistency. However, multiple studies have highlighted certain challenges, such as inadequate teacher preparedness and reliance on outdated assessments, underscoring the need for skills in critical thinking, communication, creativity, and collaboration (Al-Awaid & Hussain, 2025; Mohanna, 2024; Alrabai, 2016). This discussion answers the second research question.

C. Aligning EFL Pedagogy and Institutional Practices With Saudi Economic Transformation Goals

The implementation of EFL reforms in Saudi Arabia is essential for achieving Vision 2030, which seeks to enhance the educational system, promote economic diversification, and align teaching methods with national objectives, while fostering learner independence and engagement through innovative strategies. Multiple studies highlight important reforms in Saudi Arabia's EFL education, in line with Vision 2030, centering on curriculum innovation, human capital development, targeted teacher training programs, technology integration, research infrastructure and organizational instrumentality, greater professional autonomy, and enhanced assessment practices to boost workforce-relevant language skills and economic competitiveness (OECD, 2020; Alanazi & Curle, 2025; Alzahrani, 2025; Alali & Naif, 2020; Bhatti et al., 2022; Alothma, 2025). Curricular reform shifts focus from traditional grammar-centered teaching to communicative, task-based, and skills-oriented methodologies that better meet the evolving needs of language learners. Updated EFL curricula emphasize communicative competence, critical thinking, collaboration, and problem-solving to align with labor-market needs (Albiladi, 2022; Khawaji, 2022). Aldossary (2025) noted that integrating learning management systems (LMS) and adopting new methods, such as projects and peer assessments, alongside increased digital resources, significantly improved student satisfaction. Some Saudi universities have introduced experiential

learning programs, including capstone projects and internships, and have used alternative assessment methods such as peer assessment and reflective journals to evaluate higher-order cognitive skills and enable students to apply theoretical knowledge. Similarly, Ismail and Kassem (2022) found that Saudi EFL classrooms were adopting creative teaching approaches to enhance innovation, interaction, and practical language use, reflecting the need to adapt to a competitive, globalized economy. Alainati and Al-Hunaiyyan (2024) highlighted the need for a paradigm shift in education to cultivate crucial 21st-century skills, including critical thinking, creativity, collaboration, and digital literacy. Research on digital tools and platforms indicates that technology-mediated EFL instruction aligns well with Vision 2030's focus on innovation and digital transformation. Digital transformation is essential to Saudi Vision 2030, emphasizing economic diversification, social development, and technological advancement to promote innovation, competitiveness, and sustainable growth. Multiple studies suggest that technology-mediated EFL instruction aligns closely with Vision 2030's emphasis on innovation and digital transformation (Alamr, 2019; Al Shammre, 2025; Khreisat, 2024; Alshammari, 2015). Alfuhaid (2021) and Alsofyani and Barzanji (2024) noted that Mobile applications and AI tools, such as ChatGPT, improved speaking skills in language education by emphasizing practical language use and providing instant feedback and personalized corrections.

However, certain studies suggest a misalignment between pedagogical reforms and institutional policies with the economic transformation goals of Saudi Arabia 2030, particularly related to professional training, workload, and policy inconsistency. For example, Aldosari (2025) highlighted a misalignment with Vision 2030 goals, advocating for improved national educational outcomes, diversification of assessment methods, technology adoption, and enhanced faculty training to foster critical thinking and practical application. In a study by Alzahrani (2025), financial incentives, job security, professional development, and academic reputation were identified as key motivations for EFL instructors, helping overcome administrative challenges in pursuing Vision 2030 objectives. In a study on the challenges and complexities of teaching English in Saudi Higher Education, Alanazi (2025) highlighted the need to adapt educational methods to cultural factors, emphasizing professional development and supportive policies to enhance student engagement, motivation, and communication. Alzhvani and Alkubaidi (2020) and Al-Tamimi (2019) found that inadequate professional development and inconsistent institutional support could hinder teachers' ability to effectively implement reformed curricula. Alharbi and Saaty (2022), Alqahtani (2024), and Al-Khresheh et al. (2025) identified significant barriers to achieving Vision 2030, including insufficient feedback mechanisms, unequal gender representation, excessive strictness, instructor readiness, inadequate infrastructure, and curriculum alignment issues. Assessment reform is shifting towards performance-based evaluations to improve employability; however, a misalignment exists between teaching methods and economic goals, hindering the realization of Vision 2030 (Alghamdi et al., 2023; Afzal, 2024; Banafi, 2025). This analysis answers the third research question.

D. Sociocultural and Ideological Dimensions of EFL Policy in Saudi Economic Transformation

The EFL policy in Saudi Arabia is shaped by sociocultural and ideological factors alongside economic and social changes, affecting curriculum, teaching methods, and teacher identities. It aims to align with national goals for social progress and economic diversification. The growth of English, motivated by identity and cultural continuity, influences perceptions of EFL reforms. The literature highlights the ideological significance of English as crucial to economic development, while also noting the cultural and linguistic tensions it creates. Al-Ahdal and Al-Awaid (2018) emphasized that achieving Vision 2030 required an open-door policy to improve English proficiency and global communication skills. Their study discussed the dual objectives of EFL policy, which included modernization and economic pragmatism, along with the importance of cultural preservation. Al-Seghayer (2023) emphasized the increasing significance of English in Saudi Arabia as part of Saudi Vision 2030, particularly in education, public perception, the workforce, and media. The study highlighted English's crucial role in facilitating communication between citizens and expatriates and its importance in aligning with the country's developmental goals.

Issues of equity and gender representation in EFL materials are emphasized in the studies, highlighting a partial alignment between Vision 2030's social reform goals and existing educational practices. For example, Alqahtani (2024) examined the portrayal of Saudi women in Mega Goal EFL textbooks for secondary schools, revealing a significant gender imbalance and insufficient depiction of women in prestigious roles. The findings highlight that these textbooks do not align with Vision 2030's empowerment goals, underscoring the need for improved representation to foster gender equality. Similarly, Alkhalil (2025) explored women's language and leadership to confront the macho narrative perpetuating gender inequality. The study noted an increase in gender-neutral vocabulary while emphasizing the tension between traditional gendered language and emerging narratives that empower women in Saudi higher education, which is vital to gender policy and Vision 2030. Waterkeyn (2021) examined the sociocultural effects of English and globalization on Saudi Arabian women's higher education, highlighting the intricate relationship between cultural sensitivity, learner motivation, and the dual perception of English as both beneficial and challenging. Mills-Lampsey (2021) reported an increase in English acquisition among university-educated women in Saudi Arabia, emphasizing its positive impact on their education, empowerment, and societal roles. These studies indicate that traditional gender roles and stereotypes remain, suggesting that sociocultural change through language policy occurs gradually and unevenly.

In addition, learner-focused studies highlight affective and experiential factors that shape access to English-mediated opportunities. Almalki et al. (2023) and Sharma (2018) identified language anxiety, motivation, and confidence as key variables influencing learners' engagement with EFL, with implications for equitable participation in economically

oriented language reforms. While certain challenges remain unaddressed, studies document emerging responses to these sociocultural tensions through alternative language-planning practices. For example, Alqahtani (2022) and Alenazi (2025) emphasized the growing recognition of translanguaging and translation as educational strategies that reconcile the dominance of English with the preservation of Arabic linguistic identity, supporting economic objectives while promoting cultural and linguistic inclusivity. These findings suggest that sociocultural and ideological factors significantly influenced EFL policy outcomes under Vision 2030. While English holds economic value, its educational integration is affected by identity, equity, and evolving social norms, thereby shaping language policy at the institutional and individual levels. This analysis answers the fourth research question.

IV. CONCLUSION

This systematic review examined studies from 2013 to 2026 on language policy initiatives in Saudi Arabia, focusing on the realignment of EFL policies with Vision 2030 economic development goals, language planning for workforce readiness, the integration of EFL pedagogy with institutional practices, and the sociocultural and ideological aspects of EFL policy. It was revealed that Saudi Arabia's EFL education has made significant progress toward the goals of Vision 2030, focusing on a diversified, knowledge-based economy, enhanced global competitiveness, expanded English programs, technology integration, and recognition of English proficiency. However, it faced several challenges, including social and cultural barriers, insufficient teacher training, ineffective assessments, limited feedback, inadequate resources, and curriculum misalignment, particularly regarding the underrepresentation of females in appealing vocations. The study highlights the need for vocational and skills-based education and professional development to prepare EFL graduates for economic transformation.

ACKNOWLEDGEMENTS

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (2025/02/36868).

REFERENCES

- [1] Abbott, M. L., & Lee, K. K. (2022). Researcher-Supported Professional Learning and Development for instructors in adult L2 Programs: Factors leading to ongoing Communities of Practice. *Modern Language Journal*, 106(3), 617–634. <https://doi.org/10.1111/modl.12793>
- [2] Abdullateef, S. T., Musa Alsheikh, R., & Khalifa Ibrahim Mohammed, B. (2023). Making Saudi vision 2030 a reality through educational transformation at the university level. *Labour and Industry*, 33(2), 225–240. <https://doi.org/10.1080/10301763.2023.2184166>
- [3] Abubakar, A. A., Al-Mamary, Y. H., Singh, H. P., Singh, A., Alam, F., & Agrawal, V. (2024). Exploring factors influencing sustainable human capital development: Insights from Saudi Arabia. *Heliyon*, 10(16), e35676. <https://doi.org/10.1016/j.heliyon.2024.e35676>
- [4] Abro, A. M., Bhutto, A. R., Mughal, I., Abro, F., & Shaikh, S. S. (2025). Exploring the impact of English language proficiency on business communication effectiveness: A comprehensive research analysis. *Prosolar Insights*, 4(1), 18–27. <https://doi.org/10.62997/psi.2025a-41028>
- [5] Aburizaizah, S. J. (2022). The role of quality assurance in Saudi higher education institutions. *International Journal of Educational Research Open*, 3, 100127. <https://doi.org/10.1016/j.ijedro.2022.100127>
- [6] Agustina, V., Thamrin, N. R., & Oktoma, E. (2024). The role of English language proficiency in the global economy and business communication. *International Journal Administration Business and Organization*, 5(4), 82–90. <https://doi.org/10.61242/ijabo.24.423>
- [7] Al-Ahdal, A. A. M. H., & Al-Awaid, S. A. A. (2018). English as the lingua franca of development: finding common correlates in Saudi Arabia. *Malaysian Journal of Languages and Linguistics (MJLL)*, 7(1), 1–7.
- [8] Al-Awaid, S. A. A., & Hussain, N. M. A. (2025). English Modern Curriculum Development. *Forum for Linguistic Studies*, 7(10), 42–560. <https://doi.org/10.30564/fls.v7i10.10600>
- [9] Al Fraidan, A. A., & Alaliwi, M. (2024). Digital transformation for sustainable English language learning: Insights from Saudi Arabia and global perspectives. *Forum for Linguistic Studies*, 6(6), 439–449. <https://doi.org/10.30564/fls.v6i6.7754>
- [10] Al-Khreshah, M. H., Alshammari, S. R., & Almayez, M. (2025). Digital integration in the Saudi ELT context: a supervisory lens on teachers' technological efficacy. *Cogent Social Sciences*, 11(1). <https://doi.org/10.1080/23311886.2025.2526011>
- [11] Al Mukhallafi, T. R. (2019). English in Saudi Arabia: Status and challenges in the light of Prince Mohammad Bin Salman's Vision 2030. *International Journal of English Linguistics*, 9(4), 209–223. <https://doi.org/10.5539/ijel.v9n4p209>
- [12] Al-Mwzaiji, K. N. A., & Muhammad, A. A. S. (2023). EFL Learning and Vision 2030 in Saudi Arabia: A Critical perspective. *World Journal of English Language*, 13(2), 435. <https://doi.org/10.5430/wjel.v13n2p435>
- [13] Al-Seghayer, K. (2023). The newfound status of English in 21st-century Saudi Arabia. *International Journal of Linguistics*, 15(4), 82–103. <https://doi.org/10.5296/ijl.v15i4.21262>
- [14] Al Shammre, A. S. (2025). Digital adoption and productivity in rentier economies: Evidence from the GCC. *Systems*, 13(11), 1038. <https://doi.org/10.3390/systems13111038>
- [15] Al-Shehri, S. (2020). Transforming English language education in Saudi Arabia: Why does technology matter? *International Journal of Emerging Technologies in Learning (iJET)*, 15(6), 108–123. <https://doi.org/10.3991/ijet.v15i06.12655>

- [16] Al-Tamimi, R. (2019). Policies and issues in teaching English to Arab EFL learners: A Saudi Arabian perspective. *Arab World English Journal*, 10(2), 68–76. <https://doi.org/10.24093/awej/vol10no2.6>
- [17] Al-Zahrani, N. O. A., & Rajab, H. (2017). Attitudes and perceptions of Saudi EFL teachers in implementing Kingdom of Saudi Arabia's Vision 2030. *International Journal of English Language Education*, 5(1), 83–99. <https://doi.org/10.5296/ijele.v5i1.10733>
- [18] Afzal, N. (2024). Investigating EFL students' perceptions about the use of beyond textbook materials for university courses. *World Journal of English Language*, 14(4), 610–619. <https://doi.org/10.5430/wjel.v14n4p610>
- [19] Ahmed, M. R. A. (2024). Promoting Intercultural Pluralism: Linguistic Diversity in Saudi EFL Education. *World Journal of English Language*, 14(5), 313–321. <https://doi.org/10.5430/wjel.v14n5p313>
- [20] Alainati, S., & Al-Hunaiyyan, A. (2024). The role of educational systems in developing the twenty-first-century skills: Perspectives and initiatives of Gulf Cooperation Council countries. *IOSR Journal of Business and Management*, 26(1), 44–57. <https://doi.org/10.9790/487X-2601034457>
- [21] Alali, M., & Naif, S. (2020). *Constraints on Language Teacher Autonomy in a University Setting in Saudi Arabia* [Doctoral dissertation]. University of Southampton.
- [22] Alamr, A. S. (2019). *Digital games and English as foreign language (EFL) learning in tertiary education in Saudi Arabia* [Doctoral dissertation]. University of Wollongong.
- [23] Alanazi, K., & Curle, S. (2025). Overcoming linguistic hurdles: Challenges and strategies for English-Medium Instruction in Saudi Arabian medical education. *Social Sciences & Humanities Open*, 11, 101334. <https://doi.org/10.1016/j.ssaho.2025.101334>
- [24] Alanazi, M. J. M., & Widin, J. (2018). Exploring the role of teacher talk in Saudi EFL classroom: Importance of F-move in developing students' spoken skill. *Arab World English Journal*, 9(1), 307–320. <https://doi.org/10.24093/awej/vol9no1.22>
- [25] Alanazi, T. (2025). Addressing the Challenges and Complexities of Teaching English in Saudi Higher Education: Insights through a Sociocultural Lens. *Education Quarterly Reviews*, 8(2). <https://doi.org/10.31014/aior.1993.08.02.582>
- [26] Albiladi, W. S. (2022). English teaching and learning in Saudi Arabia: Emerging practices and strategies for achieving the educational objectives of the Saudi Vision 2030. *Journal of Language Teaching and Research*, 13(4), 781–789. <https://doi.org/10.17507/jltr.1304.11>
- [27] Aldosari, M. S. (2025). From Recall to Resilience: Reforming Assessment Practices in Saudi Theory-Based Higher Education to Advance Vision 2030. *Sustainability*, 17(21), 9415. <https://doi.org/10.3390/su17219415>
- [28] Alenazi, Y. (2025). Translation as a tool for language acquisition: Perspectives of Saudi English majors. *World Journal of English Language*, 15(3), 318–328. <https://doi.org/10.5430/wjel.v15n3p318>
- [29] Alfahadi, A. M. (2019). The impact of the implementation of Saudi Arabia's Vision 2030 on teaching English as a foreign language at public universities. *Journal of Literature, Languages and Linguistics*, 56, 32–38. <https://doi.org/10.7176/jlll/56-06>
- [30] Alfuhaid, S. R. (2021). The utilisation of Duolingo to enhance the speaking proficiency of EFL secondary school students in Saudi Arabia. *English Language Teaching*, 14(11), 9–15. <https://doi.org/10.5539/elt.v14n11p9>
- [31] Alghamdi, A., Albawardi, A., Alzuabi, N., & Alshaiji, L. (2023). Does gender matter? Motivation and learning EFL: A Saudi case study. *International Journal of Arabic-English Studies*, 23(1), 17–44. <https://doi.org/10.33806/ijaes2000.23.1.2>
- [32] Alharbi, A., & Saaty, N. (2022). Demotivating teaching practices in EFL classrooms in Saudi secondary schools. *Journal of Language Teaching and Research*, 13(6), 1324–1331. <https://doi.org/10.17507/jltr.1306.21>
- [33] Alharbi, A. S. (2024). Communicative Language Teaching Approach in a Saudi Context: A Critical Appraisal. *Eurasian Journal of Applied Linguistics*, 10(1), 60–71. <https://doi.org/10.32601/ejal.1010>
- [34] Alhazmy, E. A. (2025). Education reform and vision 2030 in Saudi Arabia: challenges and pathways. *Discover Education*, 5(1). <https://doi.org/10.1007/s44217-025-01005-4>
- [35] Alkhalil, S. F. A. (2025). Women's representations in Saudi Arabia: the case of universities' online discourses. *Cogent Arts & Humanities*, 12(1). <https://doi.org/10.1080/23311983.2025.2539718>
- [36] Allmnakrah, A., & Evers, C. (2020). The need for a fundamental shift in the Saudi education system: Implementing the Saudi Arabian economic vision 2030. *Research in Education*, 106(1), 22–40. <https://doi.org/10.1177/0034523719851534>
- [37] Almalki, R. A., Alkubaidi, M. A., & Bahanshal, D. A. (2023). EFL listening anxiety: Online and face-to-face learning. *International Journal of English Language Education*, 11(2), 35–59. <https://doi.org/10.5296/ijele.v11i2.21101>
- [38] Almesaar, O. (2024). English language in Saudi Arabia: vision 2030 in the historical prism of a clash between cultures. *Frontiers in Communication*, 9. <https://doi.org/10.3389/fcomm.2024.1205167>
- [39] Almoaily, M. (2019). Language policy and power in out-of-classroom EFL contexts: The case of English departments in Saudi Arabian higher education institutions. *International Journal of English Linguistics*, 9(2), 136–144. <https://doi.org/10.5539/ijel.v9n2p136>
- [40] Almohammadi, M. (2023). Rethinking the potentiality of Self-Efficacy Beliefs for foreign language learning in Saudi Arabia. *International Journal for Cross-Disciplinary Subjects in Education*, 14(1), 4732–4736. <https://doi.org/10.20533/ijcdse.2042.6364.2023.0582>
- [41] Alnatheer, A. A. (2013). *The role of motivation and motivational strategies in Saudi students' communicative competence in English* [Doctoral dissertation]. Queensland University of Technology.
- [42] Alnaser, A. A., Binabid, J., & Sepasgozar, S. M. E. (2024). Transforming architectural programs to meet industry 4.0 demands: SWOT analysis and insights for achieving Saudi Arabia's strategic vision. *Buildings*, 14(12), 4005. <https://doi.org/10.3390/buildings14124005>
- [43] Alneyadi, M., Aljanaahi, A., Alhammadi, E., Almarzooqi, K., & Alhammadi, K. (2023). English in Global Commerce - Understanding its Role, Benefits, and Impediments. *Research Studies in English Language Teaching and Learning*, 1(5), 245–258. <https://doi.org/10.62583/rselt.v1i5.31>
- [44] Aloairdhi, N. M., Medhesha, I. M., Almohaya, L. F., Zrair, A. N., Alharrah, B. A., Alharbi, S., Alshashaa, B. M., & Alhuqbani, M. (2025). Language acquisition and language learning among Saudi adults: A critical analysis and historical development. *English Language Teaching*, 18(9), 1–11. <https://doi.org/10.5539/elt.v18n9p1>

- [45] Alnasser, S. M. N. (2018). Investigating English language policies in Saudi higher education English departments: Staff members' beliefs. *Linguistics and Literature Studies*, 6(4), 157–168. <https://doi.org/10.13189/lls.2018.060402>
- [46] Alothma, A. T. (2025). Universities as centers for boosting human capital: A descriptive analytical study of the realization of Vision 2030 in Saudi Arabia. *Journal of Ecohumanism*, 4(2), 1585–1592.
- [47] Alrabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. *International Journal of English Linguistics*, 6(3), 21. <https://doi.org/10.5539/ijel.v6n3p21>
- [48] Alqahtani, M. H. (2024). Women in Saudi secondary school EFL textbooks: a critical study of women's empowerment as enshrined in the Saudi Vision 2030. *Frontiers in Sociology*, 9, 1307623. <https://doi.org/10.3389/fsoc.2024.1307623>
- [49] Alqahtani, M. H. (2022). The Saudi 2030 vision and translanguaging in language learning in Saudi Arabia: Looking for concord in the future. *Journal of Language and Linguistic Studies*, 18(Special Issue 1), 556–568. <https://files.eric.ed.gov/fulltext/EJ1328442.pdf>
- [50] Alqahtani, S. M. A. (2018). Teaching English in Saudi Arabia. In *English as a foreign language in Saudi Arabia* (pp. 120–137). Routledge.
- [51] Alshammari, A. K. (2015). Developing the English curriculum in the Kingdom of Saudi Arabia: Possibilities and challenges. *Arab World English Journal*, 6(4), 365–372. <https://doi.org/10.2139/ssrn.2847465>
- [52] Alsofyani, A. H., & Barzanji, A. M. (2024). The Effects of ChatGPT-Generated Feedback on Saudi EFL Learners' Writing Skills and Perception at the Tertiary Level: A Mixed-Methods Study. *Journal of Educational Computing Research*, 63(2), 431–463. <https://doi.org/10.1177/073556331241307297>
- [53] Althobaiti, A. (2025). English in Saudi Arabia: Language Policy, Sociocultural Dynamics, and the Vision 2030 Transformation. *Eurasian Journal of Applied Linguistics*, 11(3), 262–273. <https://doi.org/10.32601/ejal.11323>
- [54] Alzahrani, A. (2017). Markets and language policy in Saudi Arabia: how the English language can contribute to the success of the Saudi vision 2030. *International Journal of English Language and Linguistics Research*, 5(6), 1–12.
- [55] Alzahrani, M. A. (2025). From public schools to higher education: Exploring the motivational factors behind career transitions of EFL teachers. *Journal of International Students*, 15(8), 53–76. <https://doi.org/10.32674/mp41dn20>
- [56] Alzahrani, A. M. (2026). English as a medium of instruction and its effect on local cultural identities. *International Journal of Language and Literary Studies*, 8(1), 123–140. <https://doi.org/10.36892/ijlls.v8i1.2460>
- [57] Alzhrani, N., & Alkubaidi, M. (2020). Causes of Paradigm Shift from EFL to ESL in Higher Education in Saudi Arabia. *International Journal of Education*, 12(4), 38–48. <https://doi.org/10.5296/ije.v12i4.17652>
- [58] Asem, A., Abdullah, Y. O. R. A. G., Mohammad, A. A., & Ziyad, I. A. (2024). Navigating Digital Transformation in Alignment with Vision 2030: A Review of Organizational Strategies, Innovations, and Implications in Saudi Arabia. *Journal of Knowledge Learning and Science Technology*, 3(2), 21–29. <https://doi.org/10.60087/jklst.vol3.n2.p29>
- [59] Bahanshal, D. A. (2023). From EFL to ESL in the eye of Saudi Vision 2030. *Asian Journal of Education and Social Studies*, 48(4), 1–10. <https://doi.org/10.9734/ajess/2023/v48i41079>
- [60] Banafi, N. (2025). Status of the English language program considering Saudi Vision 2030: from students' perspectives. *Asian-Pacific Journal of Second and Foreign Language Education*, 10(1). <https://doi.org/10.1186/s40862-024-00311-y>
- [61] Barnawi, O. Z., & Al-Hawsawi, S. (2016). English Education Policy in Saudi Arabia: English Language Education Policy in the Kingdom of Saudi Arabia: Current trends, issues and challenges. In R. Kirkpatrick (Ed.), *English Language Education Policy in the Middle East and North Africa* (pp. 199–222). Springer. https://doi.org/10.1007/978-3-319-46778-8_12
- [62] Bhatti, M. A., Alyahya, M., & Alshiha, A. A. (2022). Research Culture among Higher Education Institutions of Saudi Arabia and its impact on faculty performance: Assessing the Role of Instrumentality, Research Infrastructure, and Knowledge Production. *Educational Sciences: Theory & Practice*, 22(2). <https://doi.org/10.12738/jestp.2022.2.0002>
- [63] Chowdhury, G., Ahmad, J., Koka, N. A., & Mishu, A. (2026). Bridging or Dividing? Chinese Language Integration and the Resilience of ELT in Saudi Higher Education under Vision 2030: Policy Symbolism vs. Pedagogical Realities. *World Journal of English Language*, 16(2), 309–323. <https://doi.org/10.5430/wjel.v16n2p309>
- [64] Dhawi, F., & Albaqami, T. (2017). A Survey Based Study of Strategic Directions of Saudi Universities Aligned with Economic Development Supporting Vision of 2030. *Journal of Human Resource and Sustainability Studies*, 5(3), 167–178. <https://doi.org/10.4236/jhrss.2017.53016>
- [65] Elyas, T., & Al-Hoorie, A. H. (2024). English-medium instruction in higher education in Saudi Arabia. In *The Routledge handbook of English-medium instruction in higher education* (pp. 259–271). Routledge.
- [66] Elyas, T., & Badawood, O. (2017). English Language Educational Policy in Saudi Arabia Post 21st Century: Enacted Curriculum, Identity, and Modernisation: A Critical Discourse Analysis approach. *FIRE Forum for International Research in Education*, 3(3). <https://doi.org/10.18275/fire201603031093>
- [67] Fallatah, R. (2021). *Curriculum Development in Saudi Arabia: Saudi Primary EFL Teachers' Perspectives of the Challenges of Implementing CLT into the English Curriculum in State Schools* [Thesis]. University of Exeter. <https://hdl.handle.net/10871/128108>
- [68] Faruk, S. M. G. (2013). English language teaching in Saudi Arabia: A world system perspective. *Scientific Bulletin of the Politehnica University of Timișoara, Transactions on Modern Languages*, 12(1–2), 73–80.
- [69] Fryan, L. H. A., Shomo, M. I., Hammad, R., & Imam, M. (2025). Leveraging virtual reality technology in Saudi Arabia to enhance training: A new paradigm for developing women leaders in the workplace. *Cureus*, 17(4), e81769. <https://doi.org/10.7759/cureus.81769>
- [70] Hassan, O. W. K., Mohamed, A. M., & Hamad Al-Khresheh, M. (2024). The Integration of English Language Skills in Business Education: Insights from Accounting in the Saudi Context. *Journal of Teaching in International Business*, 35(3–4), 130–158. <https://doi.org/10.1080/08975930.2025.2452529>
- [71] *Human Capability Development Program*. (2022). Available at <https://www.vision2030.gov.sa/v2030/vrps/hcdp/> (Accessed Jun 3, 2022).
- [72] Ismail, S. M., & Kassem, M. A. M. (2022). Revisiting creative teaching approach in Saudi EFL classes: Theoretical and pedagogical perspective. *World Journal of English Language*, 12(1), 142–153. <https://doi.org/10.5430/wjel.v12n1p142>

- [73] Ismail, S. M., & Namaziandost, E. (2025). Future economics and EFL policy in Saudi Arabia: A systematic review of language planning's role in digital economy transformation. *Research Journal in Advanced Humanities*, 6(2). <https://doi.org/10.58256/6r9meh15>
- [74] Jabeen, I. (2023). Implementation of new English language policy in Saudi context: perceptions, challenges, and remedies. *Journal of Language Teaching and Research*, 14(5), 1332–1343. <https://doi.org/10.17507/jltr.1405.22>
- [75] Junyi, P. (2024). Research on the Influence of Business English Proficiency on Small and Medium-Sized Enterprise Competitiveness under the Perspective of Resource-Based Theory. *Journal of Business and Economic Development*, 9(3), 69–78. <https://doi.org/10.11648/j.jbed.20240903.13>
- [76] Karim, M. R., & Hasan, M. (2024). Use of Technology in EFL Classroom: A Study to Evaluate the Training Needs of Teachers in Line with Saudi Vision 2030. *XLinguae*, 17(4), 138–152. <https://doi.org/10.18355/xl.2024.17.04.09>
- [77] Khreisat, A., Saqfahait, N., Spetan, K., & Al-Tal, R. (2024). Knowledge Economy and its Impact on Economic Growth and Per-Capita Income: Evidence from Arab Countries. *An-Najah University Journal for Research - B (Humanities)*, 38(11), 2145–2188. <https://doi.org/10.35552/0247.38.11.2300>
- [78] Khawaji, A. (2022). Transition of English language teaching in Saudi Arabia: A critical evaluative study. *Arab World English Journal*, 13(4), 265–280. <https://doi.org/10.24093/awej/vol13no4.17>
- [79] Latif, M. W. (2021). Exploring tertiary EFL practitioners' knowledge base component of assessment literacy: implications for teacher professional development. *Language Testing in Asia*, 11(1). <https://doi.org/10.1186/s40468-021-00130-9>
- [80] Latif, M. W., & Wasim, A. (2024). Investigating EFL instructors' approaches to classroom-based assessment culture: an explanatory sequential mixed-method approach. *Language Testing in Asia*, 14(1). <https://doi.org/10.1186/s40468-024-00332-x>
- [81] Mills-Lampety, S. (2021). *Cross-generational voices: Motivations and perceptions of English language learning among university-educated women of Saudi Arabia* [Ph.D. dissertation, State University of New York at Buffalo]. ProQuest Dissertations & Theses Global. Retrieved from <https://www.proquest.com/dissertations-theses/cross-generational-voices-motivations-perceptions/docview/2582078957/se-2>
- [82] Mitchell, B., & Alfuraih, A. (2017). English language teaching in the Kingdom of Saudi Arabia: past, present and beyond. *Mediterranean Journal of Social Sciences*, 8(2), 317–325. <https://doi.org/10.5901/mjss.2017.v8n2p317>
- [83] Mohanna, A. D. A. (2024). Difficulties and Challenges Encountered by Saudi EFL learners: A Diagnostic Study. *Scholars International Journal of Linguistics and Literature*, 7(10), 288–299. <https://doi.org/10.36348/sijll.2024.v07i10.002>
- [84] Mohiuddin, K., Nasr, O. A., Miladi, M. N., Fatima, H., Shahwar, S., & Naveed, Q. N. (2023). Potentialities and priorities for higher educational development in Saudi Arabia for the next decade: Critical reflections of the vision 2030 framework. *Heliyon*, 9(5), e16368. <https://doi.org/10.1016/j.heliyon.2023.e16368>
- [85] Ndiaye, A., Ullah, F., Lebbada, D., & Mohammed, E. B. (2024). English Language Teaching & Saudi Arabia Vision (2030): A Critical Scrutiny. *International Journal of Linguistics*, 16(2), 28. <https://doi.org/10.5296/ijl.v16i2.21777>
- [86] OECD. (2020). *Education in Saudi Arabia, Reviews of National Policies for Education*. OECD Publishing. <https://doi.org/10.1787/76df15a2-en>
- [87] OECD. (2019). An OECD Learning Framework 2030. In *The Future of Education and Labor Arts, research, innovation and society* (pp. 23–35). Springer. https://doi.org/10.1007/978-3-030-26068-2_3
- [88] Rauf, M., & McCallum, L. (2020). Language Assessment Literacy: Task Analysis in Saudi Universities. In *The Assessment of L2 Written English across the MENA Region* (pp. 13–41). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-53254-3_2
- [89] Saleh Alhamlan, S. A. (2025). Investigating the Academic Enablers and Hindrances of English Language Learning for Saudi EFL Secondary School Students. *American Journal of Science Education Research*. <https://doi.org/10.71010/2835-6764/ajser-e268>
- [90] Samarin, I. A., & Al-Asfour, A. A. (2023). National human resource development in transitioning societies: The case of Saudi Arabia. *New Horizons in Adult Education and Human Resource Development*, 35(1), 20–31. <https://doi.org/10.1177/19394225231171570>
- [91] Sharma, V. (2018). Influence factors in students' motivation for communicative competence in English: A case study in Saudi Arabia. *Journal of Literature, Languages and Linguistics*, 50, 37–47. <https://www.iiste.org/Journals/index.php/JLLL/article/view/41791>
- [92] Singh, H. P., Singh, A., Alam, F., & Agrawal, V. (2022). Impact of sustainable development goals on economic growth in Saudi Arabia: Role of education and training. *Sustainability*, 14(21), 14119. <https://doi.org/10.3390/su142114119>
- [93] Vision 2030. (2016). *Vision of the Kingdom of Saudi Arabia*. <https://vision2030.gov.sa/download/file/fid/417>
- [94] Waterkeyn, T. (2021). *Sociocultural impacts of English and globalisation on women in higher education in Saudi Arabia* [Doctoral dissertation, The Open University]. ProQuest Dissertations & Theses Global. Retrieved from <https://www.proquest.com/dissertations-theses/sociocultural-impacts-english-globalisation-on/docview/2543460225/se-2>
- [95] Yang, W. (2017). Tuning university undergraduates for high mobility and employability under the content and language integrated learning approach. *International Journal of Bilingual Education and Bilingualism*, 20(6), 607–624. <https://doi.org/10.1080/13670050.2015.1061474>
- [96] Yusuf, N. (2017). Changes required in Saudi universities curriculum to meet the demands of 2030 Vision. *International Journal of Economics and Finance*, 9(9), 111–116. <https://doi.org/10.5539/ijef.v9n9p111>
- [97] Zawawi, N. S. B., Aziz, S. B. B. A., Azmi, A. H. B., Razak, A. S. B. A., Nasir, N. S. B. M., & Manap, M. R. B. (2025). The impact of English proficiency in enhancing employment opportunities among university students. *International Journal of Research and Innovation in Social Science*, 9(22), 321–327. <https://doi.org/10.47772/ijriss.2025.922ileiid0031>

Saad Eid Albaqami is an Assistant Professor in the Department of Foreign Languages, Al-Baha University. He holds a PhD in Applied Linguistics from Newcastle University (UK), a Postgraduate Certificate in Research Training from Newcastle University, a

Master's degree in Linguistics from Bangor University (UK), and a Bachelor's degree in English from Taif University (Saudi Arabia). His research and teaching interests include second language writing, English vocabulary instruction, digital education, online learning, and technology-enhanced language learning.

Dhaif Alzahrani is an Assistant Professor of Applied Linguistics at King Khalid University, Saudi Arabia. His research interests include EFL pedagogy, computer-assisted language learning, AI-mediated language education, collaborative writing, and multimodal interaction analysis.

Mohammad Jamshed got his Ph.D. degree in 2018 from Aligarh Muslim University, Aligarh, UP (INDIA). His areas of interest include postcolonial literature, travel writings, comparative studies, and ESL/EFL pedagogy and research. He has presented papers at conferences and published many articles in journals of repute both in literature and ESL/ESL teaching. He is currently working as an Associate Professor, at the Department of English, College of Science & Humanities, Prince Sattam bin Abdulaziz University, Al Kharj, 11942, Kingdom of Saudi Arabia.

Mohammad Rezaul Karim is currently working as an Assistant Professor of English in the College of Science and Humanities, Prince Sattam bin Abdulaziz University, KSA. He holds a Ph.D. from Gauhati University, India. He has been teaching English language to the undergraduate students for the last 7 years. He has presented papers at both national and international conferences, published more than 25 research articles and papers in various Scopus and WOS indexed journals, and authored two books. His main area of interest is English language and comparative literature.

A DQF-MQM Evaluation of Machine Translation in English–Arabic Customs Law

Karam Damseh

School of Languages, Literacies and Translation, Universiti Sains Malaysia, Malaysia

Mozhgan Ghassemiazghandi*

School of Languages, Literacies and Translation, Universiti Sains Malaysia, Malaysia

Abstract—This study presents a corpus-based diagnostic evaluation of Google Translate's performance in translating the Revised Kyoto Convention (RKC) from English into Arabic. Building on the "Adequacy-Fluency Tradeoff" observed in recent literature, researchers employ an explanatory sequential triangulation design to assess whether high-resource NMT models can satisfy the strict compliance requirements of Customs law. The methodology applies the COMET metric to a stratified test suite of the RKC corpus (N = 1,000 segments) to establish a quantitative baseline. To diagnose potential metric discordance, researchers conducted a targeted human evaluation: a Likert-scale assessment of the highest-scoring segments (n = 100) to verify the "performance ceiling," followed by a granular DQF-MQM error analysis of the lowest-scoring segments (n = 200) conducted by a panel of Experts. The analysis identifies a profound discordance within the RKC corpus: while the system achieved a high mean COMET score of 87.70, the detailed analysis of the "performance floor" revealed 1,579 discrete errors. The resulting penalty score (1,292.5) exceeded the professional acceptance threshold for this text type by a factor of 5.7. These findings indicate that, within this specific legal corpus, Google Translate exhibits "specification-blindness," systematically failing to disambiguate polysemous "Terms of Art" required for international trade compliance. The research concludes that for the RKC, automated metrics do not yet serve as a reliable proxy for legal review, and that a Human-in-the-Loop framework remains an absolute prerequisite for the deployment of Google Translate in Customs administration.

Index Terms—compliance hazard, Google Translate, metric discordance, neural machine translation, Revised Kyoto Convention

I. INTRODUCTION

A. The Macro-Crisis: The "Compliance Hazard" in Automated Law

The rapid integration of Neural Machine Translation (NMT) into international regulatory frameworks has precipitated a "compliance hazard" that traditional evaluation paradigms fail to capture. As global trade volumes expand, Customs administrations face increasing pressure to overcome the "document lag" of manual translation through automation. While NMT has achieved remarkable milestones in linguistic fluency—often reaching "human parity" in general domains (Wang et al., 2023)—this surface-level success masks a fundamental "Adequacy-Fluency Tradeoff" (Shayegh et al., 2025). In high-stakes legal environments, improvements in neural smoothness often correlate with a degradation in semantic fidelity, creating a phenomenon where, as Elsayed (2025) observes, high fluency scores (99.2%) can coexist with critical terminology failures that 'jeopardize the machine translation (MT) performance' in high-stakes legal arguments.

For Customs administrations, this tradeoff is not merely a linguistic curiosity but a systemic risk. A translation system that prioritizes the natural flow of the target language over the rigid semantic boundaries of the source text can trigger international trade disputes, revenue leakage, and security vulnerabilities. As Elsayed (2025) demonstrates in her analysis of International Court of Justice (ICJ) proceedings, even minor linguistic deviations in legal arguments can 'deeply weaken the overall message' and misrepresent core legal logic.

This phenomenon represents what Al Maaytah (2025) describes as the 'double hand' of MT: while it enables swift linguistic interchange, it simultaneously endangers the preservation of linguistic integrity. In the legal domain, this risk is most acute in the failure to accurately render modality—the precise linguistic encoding of obligation, prohibition, and permission (Alkathery, 2023). In typologically distant language pairs like English and Arabic, this 'mediated translation process' often leads to a reduction in translation fidelity that aggregate metrics are ill-equipped to detect (Al Maaytah, 2025).

B. The Metric Conflict: The Granularity Gap

This reliability crisis is exacerbated by a persistent 'granularity gap.' While automated metrics correlate well with human judgment at the aggregate system level, recent shared tasks reveal a collapse in performance at finer resolutions. Lavie et al. (2025) report that current auto-raters 'fail to localize errors,' with the best systems rarely exceeding 35% F1

* Corresponding Author. Email: mozhgan.ghassemi@gmail.com

on span-level detection. Similarly, Zerva et al. (2024) conclude that despite strong sentence-level performance in Quality Estimation, there remains 'ample room for improvement' in fine-grained error span detection. This lack of transparency is a known limitation of widely used learned metrics, which estimate quality via a single scalar score but offer little insight into translation errors or their specific severity (Guerreiro et al., 2024). As Guerreiro et al. (2024) argue, the field is currently shifting toward more granular strategies that attempt to detail and categorize errors through span detection, bridging the gap between opaque "black-box" scores and human-interpretable diagnostics.

This "metric bias" (Kovacs et al., 2024) often rewards "reward hacking" where the model produces high-probability, fluent sequences that satisfy the metric's embedding model but fail the specific utility function of legal adequacy. Consequently, aggregate scores often obscure a "performance paradox": a translation may receive a high COMET score while containing a "Critical" error that renders the entire document legally void. This discrepancy is further evidenced by recent evaluations using professional certification standards. Zou et al. (2025) found that while COMET scores for English-Arabic translations remained uniformly high (averaging 0.97), manual assessments by certified graders revealed significant distinctions in quality, with aggregate metrics often overlooking an accumulation of low-severity errors that 'materially degrades professional-level adequacy'.

C. *The Domain Challenge: The Task-Data Mismatch in Arabic*

The challenges of NMT are further amplified in the context of Arabic, a morphologically rich language that remains low-resource in specialized domains such as finance and law (Alghamdi et al., 2024). Recent scholarship attributes the reliability gap in this language pair to a fundamental "Task-Data Mismatch" (Uhlig et al., 2025). General-purpose models like Google Translate are trained on massive, interdisciplinary web corpora that diverge significantly from the high-precision requirements of legislative drafting. This challenge is rooted in the high degree of lexical ambiguity inherent in Arabic, where homonyms and polysemes frequently lead to systemic misunderstandings in traditional MT systems (Aldawsari, 2025).

As Semenov et al. (2025) clarify, while sentence-level terminology has reached near-saturation in general domains, performance degrades precipitously in document-level contexts. This contextual decay is particularly dangerous in Customs law, where terminological consistency must be maintained across complex regulatory frameworks to ensure legal certainty. Furthermore, the translation of lexical collocations remains a significant hurdle for English-Arabic MT (Sabtan et al., 2024). Recent evaluations of Google Translate in this language pair reveal an adequacy rate of only 60% for collocations, with literal translation and omission identified as the primary causes of systemic failure (Sabtan et al., 2024). Recent evaluations confirm that even state-of-the-art NMT systems struggle to match the sophisticated requirements of Arabic's gendered grammar and syntactic flexibility (Shaalán et al., 2019; Al Maaytah, 2025). Al Maaytah (2025) identifies syntactic errors (28%) and semantic mismatches (24%) as the most prevalent challenges in this language pair, suggesting that NMT tools frequently reduce the complexity of syntax or fail to uphold the specific meanings intended by human authors.

D. *The Research Intervention: A Diagnostic Analysis of the State-of-the-Art*

To navigate this disruption, the translation industry has shifted toward structured quality management, culminating in the publication of ISO 11669:2024 and ISO 5060:2024 (Strandvik, 2025). These standards prioritize explicit "translation specifications"—such as purpose, audience, and risk tolerance—as the non-negotiable foundation for evaluation. A critical limitation of current NMT architectures, as identified by Kayano and Sugawara (2025), is the specification gap: models operate without awareness of the "translation brief" (Skopos), defaulting instead to the most probable statistical output. In the Customs domain, where the RKC serves as the "blueprint" for Customs administration of global trade (World Customs Organization [WCO], 1999), this lack of specification awareness leads to outputs that are linguistically valid but functionally disastrous.

This study addresses these critical gaps by providing a comprehensive diagnostic evaluation of Google Translate's performance in the English-Arabic Customs domain. Researchers select Google Translate not as a competitor in a comparative analysis, but as a proxy for the current State-of-the-Art (SOTA) in widely accessible NMT. Moving beyond surface-level scoring, researchers adopt a functionalist approach grounded in the 'Five Perspectives' of translation quality (Fields et al., 2014), recently reinvigorated by Strandvik (2025) to address the specific risks of automated translation. By triangulating automated COMET metrics with a granular DQF-MQM error analysis conducted by a panel of experts, the research aims to identify the specific failure points where NMT fluency masks legal liability. In doing so, this paper provides actionable insights for the integration of NMT into professional legal workflows and establishes a quantitative baseline for the "Human-in-the-Loop" (HITL) requirements necessary for autonomous deployment in specialized Customs administration.

II. LITERATURE REVIEW

A. *The Metric Evolution: From Lexical Overlap to Neural Semanticity*

The evaluation of English-Arabic MT is currently undergoing a paradigm shift from surface-level lexical metrics to deep neural architectures. Traditional metrics like BLEU have been increasingly criticized for their inability to capture the expressive integrity and contextual nuances required for morphologically rich languages (Wang et al., 2023; Gamal

et al., 2023). This inadequacy is particularly acute in Arabic, where lexical similarity does not equate to semantic adequacy (Babaali & Salem, 2022). To address this, recent scholarship has introduced "linguistic-based" evaluation methods that incorporate POS tagging and multilingual BERT models to penalize errors in sentence position and context (Beseiso et al., 2022). However, as Al-Khalifa et al. (2024) demonstrate, even advanced neural metrics like BERTScore and COMET can exhibit substantial limitations when faced with Arabic's morphological complexity, necessitating a return to human-centric, multidimensional frameworks.

To systematically uncover the weaknesses of neural metrics, recent research has turned to "challenge sets" designed to probe specific linguistic phenomena. The Translation Adequacy Challenge Set (ACES), introduced by Moghe et al. (2024), benchmarks metrics across 68 error phenomena based on the MQM ontology. Their large-scale study of 47 metrics revealed that reference-based neural metrics still rely excessively on surface-level overlap and frequently ignore information present in the source (Moghe et al., 2024). Furthermore, Moghe et al. (2024) demonstrate that LLM-based evaluators fail to provide reliable segment-level performance, often producing peaked distributions and hallucinations that render them less effective than traditional string-overlap metrics in zero-shot settings. These findings validate the need for the granular, span-based human analysis employed in this study.

B. *The "Fluency Trap" in Arabic NMT: Morpho-Syntactic Divergence*

A recurring theme in recent English-Arabic MT research is the "Fluency Trap"; a phenomenon where NMT systems produce grammatically plausible but semantically vacant outputs. Zakraoui et al. (2021) and Ameer et al. (2020) identify Arabic's complex syntax, diacritics, and long-distance dependencies as primary triggers for NMT failure. This is further evidenced by Nagi (2023), who found that when translating complex English relative clauses into Arabic, Google Translate produced a high volume of fluency errors (70.23%), suggesting that the system's ability to maintain 'linguistic well-formedness' collapses when faced with the recursive syntax of the target language. This divergence is often attributed to the 'Task-Data Mismatch' (Uhlir et al., 2025), as general-purpose models are typically trained on interdisciplinary web corpora that fail to capture the specific constraints of specialized domains (Ameer et al., 2020), which leads to "literalism" and "wrong equivalents" in colloquial or specialized registers (Sabtan et al., 2021).

Recent evaluations confirm that even state-of-the-art NMT systems struggle to match the sophisticated requirements of Arabic's gendered grammar and syntactic flexibility (Shaalan et al., 2019; Al Maaytah, 2025). Al Maaytah (2025) identifies syntactic errors (28%) and semantic mismatches (24%) as the most prevalent challenges in this language pair, suggesting that NMT tools frequently reduce the complexity of syntax or fail to uphold the specific meanings intended by human authors. This linguistic barrier is a persistent theme in recent surveys of the field. Idrys et al. (2025) observe that while significant progress has been made in neural architectures, the quality of Arabic MT systems still lags behind that of some other languages. This inherent lag necessitates a more rigorous, domain-specific evaluation to identify the linguistic and technical challenges that general-purpose systems fail to resolve in specialized contexts (Idrys et al., 2025).

C. *Fine-Grained Error Analysis: The DQF-MQM Turn*

To move beyond scalar scores, the field has pivoted toward fine-grained error typologies. The Harmonized DQF-MQM framework (Lommel, 2018) has emerged as the industry standard for diagnosing specific translation failures. Recent large-scale evaluations by Freitag et al. (2021) demonstrate that MQM-based human annotation provides a far more reliable signal of translation quality than scalar ratings. Furthermore, the framework's capacity to distinguish error severity allows for the identification of 'Critical' failures that pose safety risks (Lommel, 2018). In expressive domains, Fakih et al. (2024) report a 90% failure rate, with 'Critical' adequacy errors compromising the core message of literary texts. Conversely, in informative news contexts, Abdelaal and Alazzawie (2020) observe that while the majority of NMT errors are minor, the system persistently struggles with lexical omissions and semantic ambiguities arising from homonymy. This suggests that as the 'expressive load' of a text increases—which is the case in Customs Law—the severity of NMT failure escalates from cosmetic to critical.

For instance, in medical and literary contexts, semantic and syntactic errors remain the most frequent, requiring significant human intervention to reach professional standards (Almahasees et al., 2021; Fakih et al., 2024). This underscores a critical consensus: automated translation systems can provide a general understanding, but they cannot yet serve as a proxy for legal or technical review (Ali, 2020).

This shift is further driven by the "flattening" of meaning in NMT. A critical concern in modern linguistics is the tendency of NMT to "flatten" or oversimplify culturally and legally embedded meanings (Al Maaytah, 2025). While systems like Google Translate perform adequately on technical or literal prose where structures are predictable, they often destabilize when faced with context-dependent translations (Al Maaytah, 2025). This bifurcation of performance necessitates a mixed-methods approach that combines benchmark-based scoring with expert human review to provide a holistic evaluation of MT performance (Al Maaytah et al., 2025; Al Maaytah, 2025).

D. *The Domain-Specific Gap: From Finance to Customs Law*

While domain adaptation has shown promise in specialized fields, a significant regulatory gap remains in literature. Studies in the financial (Alghamdi et al., 2024) and industrial (Hameed et al., 2022) sectors demonstrate that fine-tuning pre-trained models on domain-specific parallel corpora can significantly outperform general-purpose engines like Google Translate. However, in the legal domain, the register-related failures of NMT remain a primary concern. Alkathery (2023)

observes that Google Translate systematically struggles with the preservation of legal discourse features, often failing to disambiguate specialized terminology. In addition, NMT systems continue to suffer from 'literal translation practices' and inadequacies in style and register, which can lead to 'semantically hollow' outputs in specialized domains (Al Maaytah, 2025). Nevertheless, as AlAfnan (2025) argues, while AI has made substantial strides, 'human oversight' remains indispensable to ensure contextual adequacy in specialized fields like law.

Despite these findings, there is a marked scarcity of empirical research addressing the Customs domain, where the intersection of international trade law under Customs administration and Arabic regulatory syntax creates a unique set of high-liability challenges that this study aims to address.

III. METHODOLOGY

A. Research Design: Explanatory Sequential Triangulation

This study employs a mixed-methods, explanatory sequential research design to evaluate the performance of Google Translate in the specialized domain of Customs law. To mitigate the biases inherent in any single evaluation instrument (Kit & Wong, 2023), the methodology achieves a triangulation perspective by integrating automated neural metrics with granular human expert analysis. This approach is designed to capture both the "performance ceiling" (optimal output) and the "performance floor" (systemic failure points), providing a detailed account of how neural fluency interacts with legal precision.

B. Corpus Selection and Stratification: The "Test Suite" Approach

The primary instrument for this study is the RKC, the foundational international legal instrument governing cross-border trade under Customs administration (WCO, 1999). Standard randomized test sets frequently mask translation deficiencies by averaging performance across disparate sentence types (Mukherjee & Shrivastava, 2023). To rigorously assess the system's competence in Customs law—a domain characterized by the complex syntactic structures found in 'Judgments' and 'Technical-writing'—researchers adopted the 'Test Suite' methodology. Following the protocol in Mukherjee and Shrivastava (2023), which prioritizes clustering by linguistic features such as Sentence Length (SL), researchers stratified the 1,000-segment parallel English-Arabic corpus (31,655 words) into three complexity tiers. This stratification also addresses the 'complexity degradation' phenomenon often observed in NMT (Koehn & Knowles, 2017; Zakraoui et al., 2021):

Short: ≤ 20 words (n=198 segments, 19.6%)

Medium: 21-40 words (n=498 segments, 49.4%)

Long: > 40 words (n=313 segments, 31.0%)

By isolating these categories, researchers move beyond monolithic scoring to perform a targeted analysis of how syntactic density impacts semantic coherence (Singh et al., 2025).

C. The Expert Panel: Ensuring Ecological Validity

To ensure high ecological validity (Läubli et al., 2018), the evaluation was conducted by a panel of three expert annotators with deep domain-specific expertise. The panel included two Senior Translators concurrently serving as Customs Officers (10+ years of experience) and one Senior Linguist specializing in Arabic (7+ years of experience). This panel represents the "professional gold standard," ensuring that the evaluation of "Critical" errors is grounded in actual regulatory risk rather than abstract linguistic preference.

Inter-Rater Reliability (IRR): To validate the consistency of the human judgments in rating the Adequacy and Fluency of the top-tier segments (n=100), IRR was calculated using Krippendorff's Alpha (α) for ordinal data. The analysis yielded moderate to substantial agreement for Adequacy ($\alpha=0.51-0.58$) and Fluency ($\alpha=0.53-0.58$), confirming that the ratings reflect a shared professional consensus rather than individual annotator bias.

D. Evaluation Framework: Decoupling Fluency From Adequacy

The study utilizes a three-tiered evaluation framework designed to prevent metric masking:

Automated Metric (COMET): Researchers deployed the Cross-lingual Optimized Metric for Evaluation of Translation (COMET) via the Machine Translation Evaluation Online (MATEO) (<https://mateo.ivdnt.org/>) platform (Vanroy et al., 2023). Unlike n-gram metrics, COMET utilizes cross-lingual embeddings to assess semantic similarity, a feature recently validated as a reliable proxy for professional judgment in multilingual settings (Uhlig et al., 2025).

Human Quality Rating (Likert Scale): A subset of high-ranked segments (n=100) was subjected to a five-point Likert scale for Adequacy and Fluency. Following Shayegh et al. (2025), these dimensions were analyzed independently to detect metric discordance.

Fine-Grained Error Analysis (DQF-MQM): The 200 lowest-ranked segments underwent an in-depth analysis using the Harmonized DQF-MQM Error Typology (Lommel, 2018). This choice is validated by Elsayed (2025), who successfully employed the DQF-MQM framework to diagnose critical terminology errors in English-Arabic legal transcripts, identifying it as a 'comprehensive and widely accepted' method for pinpointing the root causes of translation failure. As well as the work by Junczys-Dowmunt (2025) on GEMBA-MQM V2 demonstrates that MQM remains the most robust instrument for capturing fine-grained, severity-weighted terminology violations in high-stakes contexts.

E. Quality Assessment Scoring

To provide a definitive verdict on usability, a quantitative Quality Assessment was conducted using the standard TAUS DQF scoring algorithm. This methodology assigns weighted penalty points to errors based on their severity: 0.5 points for Minor errors, 1.0 point for Major errors, and 1.5 points for Critical errors. Consistent with industry benchmarks for "Business-to-Business" content, the maximum allowable penalty threshold was set at 50 points per 1,000 words. Any score exceeding this limit indicates that the translation quality is too low to be considered professionally viable without extensive re-working.

IV. RESULTS AND DISCUSSION

A. The Performance Ceiling: Quantitative and Human Validation

The initial phase of the evaluation established the "performance ceiling" of the NMT system. Quantitatively, Google Translate demonstrated robust linguistic well-formedness across the full corpus (N=1,000), achieving a mean COMET score of 87.70 (SD=5.45; 95% CI [87.4, 88.0]). This score significantly exceeds the typical baseline for raw MT output, suggesting a high degree of semantic proximity to the professional human reference.

Moghe et al. (2024) demonstrate that reference-based neural metrics frequently over-rely on surface-level lexical overlap, leading to high scores even when the translation contains hallucinations or ignores the source context. To validate that the results indicate genuine quality rather than lexical artifacts, researchers conducted a human expert analysis of the top-tier segments (n=100). The results confirmed the system's capacity for high-performance output:

Adequacy: 97.0% of segments were rated as preserving "All" (68.3%) or "Much" (28.7%) of the source meaning.

Fluency: 96.7% of segments were rated as "Flawless" (78.0%) or "Good" (18.7%).

This data confirms the existence of a 'fluency trap'; a state where, as Shayegh et al. (2025) demonstrate, metrics lean so heavily toward fluency that they fail to penalize catastrophic semantic errors. At its best, the system produces output that is statistically indistinguishable from human translation, creating a false sense of security regarding its reliability in complex contexts.

B. The Performance Floor: Fine-Grained Error Analysis

The illusion of reliability collapses when the analysis shifts to the "performance floor." The detailed analysis of the lowest-ranked segments (n=200) using the DQF-MQM framework revealed a systemic breakdown in translation integrity. The expert panel identified a total of 1,579 discrete error instances within this subset, resulting in an average Error Density of 7.895 errors per failing segment.

C. The Severity Profile and Penalty Score

The application of the TAUS DQF scoring model resulted in a cumulative penalty score of 1,292.5. Given the sample size (~4,500 words), the maximum allowable penalty threshold for professional "Business-to-Business" acceptance is 225 points. The NMT output exceeded this safety threshold by a factor of 5.7, rendering it professionally unusable without extensive post-editing.

The severity distribution (Table 1) reveals a high-risk profile:

TABLE 1
DISTRIBUTION OF ERROR SEVERITY BY HIGH-LEVEL CATEGORY

High-level Error Type	Critical	Major	Minor	Frequency (n)	Percentage (%)
Adequacy	83	278	163	524	33.20%
Terminology	80	217	161	458	29.00%
Fluency	3	149	188	340	21.60%
Style	0	30	227	257	16.20%
TOTAL	166	674	739	1,579	
Percentage (%)	10.51%	42.69%	46.80%		100.0%

Crucially, the risk is not evenly distributed. 98.2% of all Critical errors (163/166) occurred within the Adequacy and Terminology categories. This confirms that while the system rarely fails grammatically (Fluency/Style), it systematically fails to convey the correct legal meaning. This discrepancy aligns with the "Metric Bias" identified by Kovacs et al. (2024), where models optimize for high-probability fluent sequences at the expense of the specific utility function of legal adequacy.

D. Qualitative Analysis: Mechanisms of Legal Failure

The diagnostic analysis identified three specific failure modes where the NMT system's lack of specification awareness—a 'specification gap' identified by Kayano and Sugawara (2025)—created direct legal liability.

(a). Critical Over-Translation (Hallucination)

The most dangerous failure mode identified was the addition of restrictive qualifiers not present in the source text.

- Source Text: "...processing of goods for home use..."
- NMT Output: ...al-istikhdam al-manzili (Literally: "residential/household use").

- **Legal Implication:** In Customs law, "Home Use" is a Term of Art referring to goods cleared for free circulation in the domestic market (whether industrial, commercial, or personal). By adding the semantic constraint "household" (manzili), the NMT system effectively rewrote the regulation to exclude industrial importers. This hallucination fundamentally misrepresents the scope of the law, creating a false legal condition that could trigger trade disputes. This mirrors findings by Zou et al. (2025), who noted that NMT systems often stick too closely to literal meanings, failing to recognize specialized domain constraints.

(b). Critical Under-Translation (Terminological Dilution)

The system frequently defaulted to general-domain definitions for polysemous terms, stripping them of their specific regulatory force.

- **Source Text:** "Issue of stores for consumption."
- **NMT Output:** Isdar makhazin lil-istihlak (Literally: "Issuing warehouses for consumption").
- **Legal Implication:** In the context of the RKC, "stores" refers to vessel supplies (fuel, food). The NMT system failed to suppress the high-probability association of "stores" with "shops/warehouses." The resulting translation is semantically nonsensical ("consuming a warehouse"), rendering the provision functionally inoperative. This failure provides a textbook example of distributional interference (Uhlig et al., 2025), where the model defaults to the most probable general definition rather than the domain-specific definition required by the Skopos. This mirrors the 'contextual blindness' observed by Elsayed (2025), where Google Translate failed to recognize 'The Wall' as a legal entity in ICJ texts, instead rendering it as a person with an opinion. In both cases, the NMT system prioritized the most probable general meaning over the specific jurisdictional definition.

(c). Critical Omission (Jurisdictional Authority)

The analysis revealed instances where the system omitted binding provisos, altering the constitutional balance of the text.

- **Source Text:** "A Customs or Economic Union... shall, for the matters within its competence, exercise in its own name the rights..."
- **NMT Output:** (Omitted the phrase "for the matters within its competence").
- **Legal Implication:** This omission transforms a conditional delegation of power into a blanket authorization. It implies that a Customs Union has unlimited authority to act on behalf of member states, overriding national sovereignty in areas (like criminal enforcement) where competence has not been transferred. While the Arabic output was grammatically fluent, it was legally unconstitutional. Where terminological consistency must be maintained across complex regulatory frameworks to ensure the precise expression of legal modality; the linguistic encoding of obligation and permission (Alkatheery, 2023). As Elsayed (2025) argues, the failure of NMT to preserve these 'deontic' features in international legal discourse can 'deeply weaken the overall message' and lead to the misinterpretation of core regulatory logic.

E. The Complexity Penalty: Syntactic Degradation

Finally, the stratified analysis confirms that syntactic complexity is a primary trigger for these failures. While "Long" sentences (>40 words) constituted only 19.6% of the original corpus, they accounted for 33.0% of the total error sample.

This complexity degradation manifests as a loss of contextual integrity. As sentence length increases, the NMT attention mechanism struggles to maintain long-distance dependencies, leading to the decoupling of conditional clauses (e.g., "provided that...") from their subjects. This data validates the hypothesis by Al Maaytah (2025) that current NMT architectures, despite their fluency, lack the cognitive depth required to parse the recursive syntax characteristic of international legal conventions.

V. CONCLUSION

This study provides an in-depth diagnostic of Google Translate's performance in the English-Arabic Customs domain, uncovering a profound safety gap that aggregate metrics fail to register. While the system demonstrates a high "performance ceiling" in general contexts, its "performance floor" in complex regulatory segments is characterized by systemic failures in adequacy and terminology. The research concludes that NMT, in its current state, cannot function as an autonomous agent for the translation of international Customs conventions. The 5.7x failure gap identified in this research provides a quantitative baseline for the compliance hazard inherent in unmonitored NMT deployment.

The results underscore a persistent performance paradox: high fluency scores actively camouflage critical adequacy failures. This "metric discordance" suggests that current evaluation paradigms—and the NMT objectives themselves—cannot reliably resolve the tradeoff in favor of strict legal precision. The findings confirm that while NMT has mastered the form of legal Arabic, it remains blind to the function of legal obligation.

To address these limitations, future research must focus on the development of Grounded Metrics such as COMET-poly-ic (Züfle et al., 2025), which retrieve human-scored legal segments to calibrate automated scores. Additionally, the use of granular post-edits as 'preference data' for fine-tuning models via Direct Preference Optimization (DPO) (Uhlig et al., 2025) offers a promising pathway toward aligning NMT systems with the rigorous demands of international law.

Furthermore, integrating human error markings directly into the inference process, as demonstrated by Berger et al. (2024), could enable real-time self-correction for critical legal terminology. Future research should specifically evaluate the efficacy of xCOMET-XXL (Guerreiro et al., 2024) in the legal domain. By utilizing a metric that integrates fine-grained error detection directly into the scoring process, Customs administrations may be able to automate the identification of localized critical errors and hallucinations, thereby reducing the cognitive load on human posteditors while maintaining regulatory safety.

REFERENCES

- [1] Abdelaal, N. M., & Alazzawie, A. (2020). Machine translation: the case of Arabic-English translation of news texts. *Theory and Practice in Language Studies*, 10(4), 408–418. <https://doi.org/10.17507/tpls.1004.09>
- [2] AlAfnan, M. A. (2025). Artificial Intelligence and Language: Bridging Arabic and English with Technology. *Journal of Ecohumanism*, 3(8). <https://doi.org/10.62754/joe.v3i8.4961>
- [3] Aldawsari, H. A. H. (2025). Evaluating the Performance of Large Language Models on Arabic Lexical Ambiguities: A Comparative Study with Traditional Machine Translation Systems. *World Journal of English Language*, 15(3), 354. <https://doi.org/10.5430/wjel.v15n3p354>
- [4] Alghamdi, E. A., Zakraoui, J., & Abanmy, F. A. (2024). Domain adaptation for Arabic machine Translation: Financial Texts as a case study. *Applied Sciences*, 14(16), 7088. <https://doi.org/10.3390/app14167088>
- [5] Ali, M. A. (2020). Quality and Machine Translation: An Evaluation of Online Machine Translation of English into Arabic Texts. *Open Journal of Modern Linguistics*, 10(05), 524–548. <https://doi.org/10.4236/ojml.2020.105030>
- [6] Alkathery, E. R. (2023). Google Translate Errors in Legal Texts: Machine Translation Quality Assessment. *Arab World English Journal for Translation and Literary Studies*, 7(1), 208–219. <https://doi.org/10.24093/awejtls/vol7no1.16>
- [7] Al-Khalifa, H., Al-Khalefah, K., & Haroon, H. (2024). Error analysis of Pretrained Language Models (PLMS) in English-to-Arabic machine translation. *Human-Centric Intelligent Systems*, 4(2), 206–219. <https://doi.org/10.1007/s44230-024-00061-7>
- [8] Al Maaytah, S. A. A. (2025). Evaluating Three neural machine translation platforms for English-Arabic translation: A Comparative study of linguistic accuracy and Cultural fidelity. *World Journal of English Language*, 16(2), 1. <https://doi.org/10.5430/wjel.v16n2p1>
- [9] Al Maaytah, S. A., Aalzobidy, S. A., & Alwidyan, M. F. (2025). Using machine translation English - Arabic procedures and challenges - A systematic review. *Power System Technology*, 49(1), 588–607. Retrieved February 18, 2026, from <https://powertechjournal.com/index.php/journal/article/view/1582>
- [10] Almahasees, Z., Meqdadi, S., & Albudairi, Y. (2021). Evaluation of google translate in rendering English COVID-19 texts into Arabic. *Journal of Language and Linguistic Studies*, 17(4), 2065–2080. <https://doi.org/10.52462/jlls.149>
- [11] Ameer, M. S. H., Meziane, F., & Guessoum, A. (2020). Arabic Machine Translation: A survey of the latest trends and challenges. *Computer Science Review*, 38, 100305. <https://doi.org/10.1016/j.cosrev.2020.100305>
- [12] Babaali, B., & Salem, M. (2022). Arabic machine Translation: A panoramic survey. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4312742>
- [13] Berger, N., Riezler, S., Exel, M., & Huck, M. (2024). Prompting Large Language Models with Human Error Markings for Self-Correcting Machine Translation. In *Proceedings of the 25th Annual Conference of the European Association for Machine Translation* (Vol. 1, pp. 636–646). European Association for Machine Translation (EAMT). Retrieved February 18, 2026, from <https://aclanthology.org/2024.eamt-1.54/>
- [14] Beseiso, M., Tripathi, S., Al-Shboul, B., & Aljadid, R. (2022). Semantics based English-Arabic machine translation evaluation. *Indonesian Journal of Electrical Engineering and Computer Science*, 27(1), 189–197. <https://doi.org/10.11591/ijeecs.v27.i1.pp189-197>
- [15] Elsayed, A. S. O. (2025). When machines meet gavel: a case study of the English–Arabic machine translation of the Egyptian arguments before the International Court of Justice (2024). *Language and Semiotic Studies*, 11(4), 661–690. <https://doi.org/10.1515/lass-2025-0054>
- [16] Fakihi, A., Ghassemiazghandi, M., Fakihi, A. H., & Singh, M. K. M. (2024). Evaluation of Instagram’s neural machine translation for literary texts: An MQM-Based Analysis. *GEMA Online Journal of Language Studies*, 24(1), 213–233. <https://doi.org/10.17576/gema-2024-2401-13>
- [17] Fields, P., Hague, D. R., Kobayashi, G. S., Lommel, A., & Melby, A. (2014). What is quality? A management discipline and the translation industry get acquainted. *Tradumàtica Tecnologies De La Traducció*, 12, 404–412. <https://doi.org/10.5565/rev/tradumatica.75>
- [18] Freitag, M., Foster, G., Grangier, D., Ratnakar, V., Tan, Q., & Macherey, W. (2021). Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation. *Transactions of the Association for Computational Linguistics*, 9, 1460–1474. https://doi.org/10.1162/tacl_a_00437
- [19] Gamal, D., Alfonse, M., Jiménez-Zafra, S. M., & Aref, M. (2023). Case study of improving English-Arabic translation using the Transformer Model. *International Journal of Intelligent Computing and Information Sciences*, 23(2), 105–115. <https://doi.org/10.21608/ijicis.2023.210435.1270>
- [20] Guerreiro, N. M., Rei, R., Van Stigt, D., Coheur, L., Colombo, P., & Martins, A. F. T. (2024). xCOMET: Transparent Machine Translation Evaluation through Fine-grained Error Detection. *Transactions of the Association for Computational Linguistics*, 12, 979–995. https://doi.org/10.1162/tacl_a_00683
- [21] Hameed, D. A., Faisal, T. A., Abbas, A. K., Ali, H. A., & Hasan, G. T. (2022). DIA-English-Arabic neural machine translation domain: sulfur industry. *Indonesian Journal of Electrical Engineering and Computer Science*, 27(3), 1619–1624. <https://doi.org/10.11591/ijeecs.v27.i3.pp1619-1624>
- [22] Idrys, F. Z. E., Hourri, S., Miqdadi, I. E., Hayati, A., Namir, Y., Ncir, B., & Kharroubi, J. (2025). Unlocking the language barrier: A Journey through Arabic machine translation. *Multimedia Tools and Applications*, 84(14), 14071–14104. <https://doi.org/10.1007/s11042-024-19551-8>

- [23] Junczys-Dowmunt, M. (2025). GEMBA V2: Ten Judgments Are Better Than One. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 926–933). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.wmt-1.67>
- [24] Kayano, Y., & Sugawara, S. (2025). Specification-Aware Machine Translation and Evaluation for Purpose Alignment. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 113–141). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.wmt-1.7>
- [25] Kit, C., & Wong, B. T. M. (2023). Evaluation in Machine Translation and Computer-Aided Translation. In *Routledge Encyclopedia of Translation Technology* (2nd edition, pp. 219–244). Routledge. <https://doi.org/10.4324/9781003168348-13>
- [26] Koehn, P., & Knowles, R. (2017). Six challenges for neural machine Translation. In *Proceedings of the First Workshop on Neural Machine Translation* (pp. 28–39). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-3204>
- [27] Kovacs, G., Deutsch, D., & Freitag, M. (2024). Mitigating Metric Bias in Minimum Bayes Risk Decoding. In *Proceedings of the Ninth Conference on Machine Translation* (pp. 1063–1094). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.wmt-1.109>
- [28] Läubli, S., Sennrich, R., & Volk, M. (2018). Has Machine Translation Achieved Human Parity? A Case for Document-level Evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 4791–4796). Association for Computational Linguistics. <https://doi.org/10.18653/v1/d18-1512>
- [29] Lavie, A., Hanneman, G., Agrawal, S., Kanojia, D., Lo, C., Zouhar, V., Blain, F., Zerva, C., Avramidis, E., Deoghare, S., Sindhuja, A., Wang, J., Adelani, D. I., Thompson, B., Kocmi, T., Freitag, M., & Deutsch, D. (2025). Findings of the WMT25 Shared Task on Automated Translation Evaluation Systems: Linguistic Diversity is Challenging and References Still Help. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 436–483). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.wmt-1.24>
- [30] Lommel, A. (2018). Metrics for Translation Quality Assessment: A case for Standardising error typologies. In *Machine translation* (pp. 109–127). Springer. https://doi.org/10.1007/978-3-319-91241-7_6
- [31] Moghe, N., Fazla, A., Amrhein, C., Kocmi, T., Steedman, M., Birch, A., Sennrich, R., & Guillou, L. (2024). Machine Translation Meta Evaluation through Translation Accuracy Challenge Sets. *Computational Linguistics*, 51(1), 73–137. https://doi.org/10.1162/coli_a_00537
- [32] Mukherjee, A., & Shrivastava, M. (2023). IIIT HYD's Submission for WMT23 Test-suite Task. In *Proceedings of the Eighth Conference on Machine Translation* (pp. 246–251). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.wmt-1.24>
- [33] Nagi, K. A. (2023). Arabic and English relative clauses and machine translation challenges. *Mağalla' Al-dirāsāt Al-iğtimā' iyya'*, 29(3), 145–165. <https://doi.org/10.20428/jss.v29i3.2180>
- [34] Sabtan, Y. M. N., Hussein, M. S. M., Ethelb, H., & Omar, A. (2021). An evaluation of the accuracy of the machine translation systems of social media language. *International Journal of Advanced Computer Science and Applications*, 12(7). <https://doi.org/10.14569/ijacsa.2021.0120746>
- [35] Sabtan, Y. M. N., Omar, A., & Hamouda, W. I. (2024). Exploring the Role of Machine Translation in Translating English Collocations into Arabic: Insights from Student Translators. *World Journal of English Language*, 14(2), 74. <https://doi.org/10.5430/wjel.v14n2p74>
- [36] Semenov, K., Huang, X., Zouhar, V., Berger, N., Zhu, D., Oncevay, A., & Chen, P. (2025). Findings of the WMT25 Terminology Translation Task: Terminology is Useful Especially for Good MTs. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 554–576). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.wmt-1.30>
- [37] Shaalan, K., Siddiqui, S., Alkhatib, M., & Monem, A. A. (2019). Challenges in Arabic natural language processing. In *Computational Linguistics, Speech and Image Processing for Arabic Language* (pp. 59–83). World Scientific. https://doi.org/10.1142/9789813229396_0003
- [38] Shayegh, B., Peter, J., Vilar, D., Domhan, T., Juraska, J., Freitag, M., & Mou, L. (2025). Feeding Two Birds or Favoring One? Adequacy–Fluency Tradeoffs in Evaluation and Meta-Evaluation of Machine Translation. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 269–285). <https://doi.org/10.18653/v1/2025.wmt-1.16>
- [39] Singh, K. B., Kumar, D., & Ekbal, A. (2025). Evaluation of LLM for English to Hindi Legal Domain Machine Translation Systems. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 823–833). <https://doi.org/10.18653/v1/2025.wmt-1.57>
- [40] Strandvik, I. (2025). Translation quality and the role of specifications – How standards can help the translation sector today. *Across Languages and Cultures*, 26(S), 5–24. <https://doi.org/10.1556/084.2025.01057>
- [41] Uhlig, K., Wuebker, J., Reinauer, R., & Denaro, J. (2025). Cross-lingual Human-Preference Alignment for Neural Machine Translation with Direct Quality Optimization. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 31–51). <https://doi.org/10.18653/v1/2025.wmt-1.2>
- [42] Vanroy, B., Tezcan, A., & Macken, L. (2023). MATEO: MT evaluation Online. In *The 24th Annual Conference of the European Association for Machine Translation (EAMT 2023)* (pp. 499–500). European Association for Machine Translation. Retrieved February 18, 2026, from <https://aclanthology.org/2023.eamt-1.52/>
- [43] Wang, L., Lyu, C., Ji, T., Zhang, Z., Yu, D., Shi, S., & Tu, Z. (2023). Document-Level Machine Translation with Large Language Models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. <https://doi.org/10.18653/v1/2023.emnlp-main.1036>
- [44] World Customs Organization. (1999). *International convention on the simplification and harmonization of customs procedures (as amended)*. Brussels: World Customs Organization. Retrieved February 18, 2026, from https://www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/conventions/kyoto-convention/revised-kyoto-convention/body_gen-annex-and-specific-annexes.pdf?la=en
- [45] Zakraoui, J., Saleh, M., Al-Maadeed, S., & Alja'am, J. M. (2021). Arabic Machine Translation: A survey with challenges and future directions. *IEEE Access*, 9, 161445–161468. <https://doi.org/10.1109/access.2021.3132488>
- [46] Zerva, C., Blain, F., De Souza, J. G. C., Kanojia, D., Deoghare, S., Guerreiro, N. M., Attanasio, G., Rei, R., Orasan, C., Negri, M., Turchi, M., Chatterjee, R., Bhattacharyya, P., Freitag, M., & Martins, A. (2024). Findings of the Quality Estimation Shared

- Task at WMT 2024: Are LLMs Closing the Gap in QE? In *Proceedings of the Ninth Conference on Machine Translation* (pp. 82–109). <https://doi.org/10.18653/v1/2024.wmt-1.3>
- [47] Zou, L., Saeedi, A., & Koby, G. S. (2025). Beyond automated metrics: Assessing GPT-4o and Google Translate against professional translation standards. *SKASE Journal of Translation and Interpretation*, 18(2). <https://doi.org/10.33542/jti2025-s-9>
- [48] Züfle, M., Zouhar, V., Dinh, T. A., Polo, F. M., Niehues, J., & Sachan, M. (2025). COMET-poly: Machine Translation Metric Grounded in Other Candidates. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 887–904). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.wmt-1.63>

Karam Damseh is a Ph.D. candidate at the School of Languages, Literacies, and Translation, Universiti Sains Malaysia. His research interests focus on Machine Translation Evaluation. With more than seven years of experience in Customs administration, his research investigates the intersection of computational linguistics and legal compliance, aiming to mitigate liability risks in automated translation.

Mozhgan Ghassemiazghandi, Ph.D., is a Senior Lecturer at the School of Languages, Literacies and Translation, Universiti Sains Malaysia. Her research interests focus on Translation Technology, Machine Translation, and Audiovisual Translation. In addition to her academic work, she is a professional translator and subtitler with more than a decade of industry experience, combining theoretical insight with practical expertise in translation practice and technology.

Authorial Ideology and Worldview in Kijne's *Kota Emas*: A Genetic Structuralist Analysis

Insum Malawat

Faculty of Teacher Training and Education, University of Papua, Manokwari, Indonesia

Hengki

Faculty of Teacher Training and Education, University of Papua, Manokwari, Indonesia

Burhanuddin Arafah

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Abstract—This study investigates the authorial ideology and worldview in Izak Samuel Kijne's short story *Kota Emas* by employing Lucien Goldmann's genetic structuralism theory. The analysis focuses on five key components: human facts, collective subjects, worldviews, narrative structures, and the dialectic of understanding and explanation. The findings indicate that individual human emotions—anger, remorse, sorrow, and joy—reflect the power relations between Regina and Tomi, as well as the moral transformation experienced by both characters. Social facts emerge through depictions of communal cooperation, Wondama cultural practices, and mystical elements that represent the local belief systems. The collective subject illustrates the tension between Western and Eastern value systems, which is ultimately reconciled through friendship. Kijne's worldview is expressed through messages emphasizing equality, compassion, and human solidarity. Furthermore, the narrative's structural opposition between imagination and reality reinforces Kijne's vision of Papua as a harmonious and dignified "golden city".

Index Terms—genetic structuralism, authorial ideology, worldview, Papuan, *Kota Emas*

I. INTRODUCTION

Modern literary studies no longer view literary works as separate from their sociocultural environments (Arafah et al., 2024; Arifuddin et al., 2024). Instead, literature is seen as a cultural product intertwined with the social, historical, and ideological structures of its society (Arafah & Kaharuddin, 2019; Lestari et al., 2025). This perspective, emphasizing the link between textual representation and social reality, is crucial in Indonesian literature studies (Purwaningsih et al., 2020; Suhadi et al., 2022), especially from regions with colonial histories, like Papua. Here, literary texts often portray individual experiences while expressing a collective consciousness shaped by power relations, cultural transformation, and shifting identities (Fadillah et al., 2022; Rano et al., 2026; Yudith et al., 2023). Thus, the novel serves as a medium reflecting cultural values, social tensions, and human experiences (Arafah, 2018; Arafah et al., 2024; Arafah et al., 2026; Hasyim & Arafah, 2023b). This aligns with studies showing that cultural products often embed stereotypes and value systems reproducing broader ideological structures (Taqdir, 2025).

In addition, numerous literary and cultural studies have demonstrated how texts may function as mirrors of specific social conditions, such as gendered representations within colonial settings (Afiah et al., 2022; Asriyanti et al., 2022) or critiques of cultural and environmental inequalities (Jaelani et al., 2024; Siwi et al., 2022). Environmental inequalities include species loss, ecosystem damage, and the fading of the concept of cultural heritage (Manugeren et al., 2023; Takwa et al., 2024a). Similar tendencies are visible in research on local cultural traditions: studies on ritual discourse, for instance, reveal that traditional texts encode belief systems, collective norms, and the worldviews of the communities that produce them (Effendy et al., 2025; Takwa et al., 2024b; Takwa et al., 2025; Takwa et al., 2026). Due to changes in lifestyle, modernization, and technological development, the locality tends to be eroded (Arafah et al., 2025; Arifin et al., 2022). The existence of technological-based media has shifted human behavior into a massive user of internet or social media and the users keep growing in number in a stage where digital technologies are changing how people behave (Arafah & Hasyim, 2023a; Arafah & Hasyim, 2023b; Hasjim et al., 2026; Hasyim & Arafah, 2023a). On a broader level, issues of ideology and representation in texts can also be observed in analyses of how gender, identity, and power relations are symbolically constructed, as reflected in research on gender representation in foreign language textbooks (Taqdir et al., 2025; Taqdir et al., 2026). Discrimination against women is also a common issue presented in literary works (Asri et al., 2023). Furthermore, these conditions are portrayed in literary works as an attempt to increase people's awareness of the current situation using language as the communication tool in which language and culture has an interdependent relation (Arafah et al., 2023a). Literary works are also introduced in academic field as a source of material, provided both in printed materials as well as in a form of electronic (Arafah et al., 2023b). As a result of globalization era, literary works as the source of materials are sometimes provided in English in which the students need to learn in order to understand

the materials given (Kaharuddin et al., 2023; Kuswanty et al., 2023). Therefore, incorporating digital media and digital literacy in this digital age is beneficial as long as it meets the needs of both students and teachers (Idhan et al., 2026; Kaharuddin et al., 2024; Usman et al., 2024).

Literary work serves as a societal purpose as it reflects social reality (Arafah et al., 2023c). Lucien Goldmann's genetic structuralism provides a theoretical foundation for examining these connections between text and society. Goldmann argues that literary works represent the worldview of a collective subject expressed through narrative structure, symbolic oppositions, character relationships, and thematic development. Accordingly, literature is not merely the outcome of an individual's imagination but the product of a dialectical relationship between the author's consciousness and the broader social-historical structure in which the text is embedded. This view aligns with the argument that literary texts are always shaped by social structures and relations of power, including dominant ideologies that operate implicitly through narrative forms (Jamiluddin et al., 2026; Jusdalyana et al., 2024; Yudith et al., 2024). Recent scholarship also illustrates how social structures and value systems are mediated through language, symbolism, and character interactions in literary works (Radjaban et al., 2025; Sunyoto et al., 2022). The language style such as symbols and metaphors is a way to put some artistic values to convey ideas beautifully and effectively (Baa et al., 2023). However, since it needs to be observed by its context, there can be an issue of how the language is not understood properly (Iksora et al., 2022).

Izak Samuel Kijne's short story *Kota Emas* represents a particularly compelling text for such an analysis. Written in 1958, the story intertwines a simple tale of childhood friendship with deeper ideological tensions between Papuan society and external forces of modernity, education, and colonial structures of power. The characters Regina and Tomi function not only as narrative figures but also as symbolic embodiments of two contrasting value systems—the West and the East, modernity and tradition, superiority and subordination—that inform the interpretive dynamics of the story (Mare et al., 2025). From a theoretical standpoint, these tensions resonate with the ideological mechanisms frequently found in literary narratives, including domination, symbolic opposition, and the negotiation of power (Arafah et al., 2023e; Halil et al., 2024). Furthermore, literary texts often reveal the complexities of human experiences within broader social structures (Mutmainnah et al., 2022; Yudith et al., 2023).

Although *Kota Emas* has been widely used for cultural education and moral instruction, scholarly work employing genetic structuralism to examine this narrative remains limited. Studies of Papuan literature typically focus on issues of ethnic identity, land conflict, cultural resistance, and state-indigenous power dynamics. These perspectives are valuable but seldom address how collective consciousness and the author's worldview systematically operate within the textual structure. However, following Goldmann, literary works must be read as structures homologous to the social structures that produce them—an approach increasingly adopted in studies of ideology and representation in literature (Mokoginta & Arafah, 2021; Arnawa & Arafah, 2023).

Kijne's biographical and historical context further justifies the use of this approach. During nearly two decades of service in Papua as an educator and missionary, Kijne witnessed firsthand profound social changes, inequalities between settlers and indigenous communities, and challenges related to human development in the region. These experiences shaped the characters, symbols, conflicts, and themes in *Kota Emas*. The motif of the 'golden city,' for example, may be interpreted as a projection of collective aspirations rooted not only in moral imagination but also in local knowledge systems that sustain dignity, continuity, and communal resilience (Takwa et al., 2026).

As one of the earliest narrative representations of Papua in Indonesian literature, *Kota Emas* reflects global dynamics such as colonial encounters, cultural negotiation, and identity search. Thus, applying Goldmann's framework to this text fills a gap in Papuan literary studies and extends genetic structuralism to Eastern Indonesian literary contexts, under-represented in international scholarship. This study seeks to answer: How do the narrative structure and character representations in Izak Samuel Kijne's *Kota Emas* reflect collective consciousness and authorial ideology within Papua's socio-historical context? The findings contribute to (1) enriching Papuan literary scholarship by linking textual analysis to social structure and (2) expanding Goldmann's genetic structuralism in Indonesian literary studies, particularly from geographically and politically marginalized regions.

II. LITERATURE REVIEW

A. Lucien Goldmann's Genetic Structuralism

Genetic structuralism integrates the internal structure of a literary work with the socio-historical conditions that shape its production. According to Lucien Goldmann, literature is not merely the outcome of individual imagination but the expression of a collective consciousness belonging to a particular social group situated within specific historical conditions. This perspective contrasts with formalist structuralism, which primarily focuses on the internal relations among textual elements. By comparison, genetic structuralism emphasizes that literary structures emerge from historical, economic, and ideological dynamics that inform the consciousness of their creator (Goldmann, 1977).

From this viewpoint, a literary work maintains a homologous relationship with social structure; that is, the internal relations within the text correspond to relations within social reality. Thus, the aim of literary analysis is not only to identify narrative or symbolic patterns but also to uncover the worldview articulated through the text as a manifestation of the collective socio-historical experience.

- (a) Human Facts
- (b) Collective Subject

- (c) Worldview
- (d) Literary Structure
- (e) Dialectics of Understanding and Explanation

B. Previous Studies on Genetic Structuralism

In Indonesia, genetic structuralism has been applied in several literary studies, particularly those examining popular novels or texts that foreground social conflict. Pratiwi et al. (2017) employed this approach to investigate the worldview represented in *Perahu Kertas* by Dewi Lestari, while Sembada and Andalas (2019) used the same theoretical framework to uncover the social realities depicted in *Laut Bercerita*. Another study by Rukiyah (2019) explored the moral value of responsibility in *Laskar Pelangi* through the lens of genetic structuralism.

Although these studies contribute to the development of Goldmann's theoretical application in Indonesian literary criticism, most remain centered on texts set in urban contexts or reflect Java-centric cultural landscapes. The use of genetic structuralism to analyze literature from Eastern Indonesia, particularly Papua, remains limited. This gap highlights the need for further exploration of literary works originating from regions with complex colonial histories and distinct sociocultural dynamics.

Moreover, prior research tends to be descriptive in nature and has not fully integrated an in-depth analysis of narrative structure with the historical conditions and the ideological perspectives of the author. This methodological limitation presents a clear opportunity for a more comprehensive study of *Kota Emas*, especially in uncovering the author's worldview within Papua's socio-historical setting.

C. Studies on Papuan Literature and Cultural Representation

Recent scholarship on Papuan literature demonstrates that works from this region frequently foreground themes such as identity, resistance, land conflict, colonial encounters, and Indigenous communities' struggles. Analyses of Papuan poetry, for example, often reveal strong critiques of social and political marginalization. Studies of Papuan short stories and novels tend to focus on the representation of local cultural values, human–nature relationships, and narratives of resistance against state power and external settlers.

Despite these contributions, much of the existing research has not examined the ideological structures or the collective worldview embedded in these texts through systematic theoretical frameworks such as genetic structuralism. Papuan literature is often approached as an anthropological record rather than an expression of collective consciousness that can be interpreted structurally and dialectically.

In this context, *Kota Emas* occupies a distinctive position. The short story not only portrays Papuan cultural elements but also encapsulates symbolic representations of power relations, value conflicts, and social aspirations—elements that are particularly amenable to analysis through Goldmann's theory. Consequently, this study offers a new perspective by reading Papuan literature as an articulation of a group's worldview rather than merely as documentation of cultural practices.

III. METHOD

This study adopted a descriptive qualitative approach supported by content analysis. This approach is appropriate for examining literary texts as representations of social and ideological phenomena that cannot be quantified or measured. Content analysis was employed to identify, classify, and interpret meaningful units in *Kota Emas*, within the framework of Lucien Goldmann's genetic structuralism.

The primary data source is the short story *Kota Emas* by Izaak Samuel Kijne, published by J.B. Wolters in 1958. The text is regarded as an external official document because it was formally published and intended for public reading. Data were collected through close reading and systematic note taking. The text was read intensively to locate relevant elements, which were then recorded and organized into analytical tables corresponding to Goldmann's theoretical components.

Data analysis was performed in several stages. First, the data were classified according to the five major components of genetic structuralism: (1) human facts, (2) collective subject, (3) worldview, (4) literary structure, and (5) dialectic of understanding and explanation. Second, each category was examined using content analysis to identify patterns of values, conflicts, cultural oppositions, and social relations that emerged from the narrative. Third, a dialectical analysis was conducted by relating individual textual fragments to the overall narrative structure to establish homologies between the story and the broader social realities of Papuan society.

The analytic procedure follows Mahsun's (2012) qualitative analysis model, consisting of: (1) data provision through intensive reading, (2) data reduction through selection and categorization, (3) data presentation in the form of tables and analytical narratives, and (4) conclusion drawing by examining the dialectical relationship between the textual data and the collective consciousness represented in the story. The findings are presented through interpretive descriptions that link textual elements to the socio-historical context of Papua.

IV. RESULT AND DISCUSSION

A. *Individual Human Facts: Emotions and Character Transformation*(a). *Expressions of Anger*

The following excerpt illustrates Regina's outburst of anger toward Tomi after the accidental damage to her doll, Sarina: ...With a smile, she watched the children playing. Regi and Tom were chasing each other... Tomi jumped forward. Alas, he did not see the doll. His foot struck the chair. Tom fell. The chair fell. The doll fell. Oh no! Sarina's head was broken! Regina was furious. Angry and distressed. She sobbed uncontrollably. She ran toward Tomi. Tom trembled, his face pale. It was a disaster! Regi slapped Tom, shouting: "You stupid boy! Why weren't you careful?" "Oh, oh, my doll is completely ruined." Tom could not say anything. He was deeply ashamed... (Kijne, 1958, p. 11).

This passage depicts Regina's intense anger toward Tomi, revealing more than frustration over a broken toy. The incident highlights an embedded power structure in their relationship. Regina assumes the role of a "little master," feeling entitled to scold, strike, and blame, while Tomi, an Indigenous child, lacks authority or confidence to defend himself.

Regina's response reflects internalized colonial values. Her quickness to assert superiority—through insults and aggression—demonstrates how hierarchies of white–nonwhite, outsider–native, and master–servant shape interactions. The harsh reprimand, coupled with Tomi's wordless shame, underscores how inequality becomes internalized by the marginalized.

Tomi's reaction—silence, fear, and embarrassment—indicates how colonial power molds consciousness. He doesn't merely feel regret; he accepts blame as if inherently at fault. This moment becomes an entry point for reading cultural tension, social disparity, and domination ideology in the colonial setting of Papua.

Narratively, this scene initiates a trajectory of emotional and moral transformation. The anger rupturing their relationship serves as a catalyst for subsequent emotions—remorse, sorrow, and reconciliation. This progression highlights the characters' moral development and reinforces the narrative's central themes: friendship, equality, and dismantling hierarchical boundaries.

(b). *Expressions of Guilt*

The following passage illustrates a moment of profound self-awareness and remorse experienced by Regina after realizing that her harsh behavior has deeply hurt Tomi:

Oh, Regi! Her mouth fell open, her eyes widened, and for a moment she stood frozen. Then she turned back slowly, sobbing uncontrollably. "Oh, it is my fault. I am a wicked child. I was so angry at Tomi." As Regi descended from the mountain, many thoughts crowded her mind. "Good Tom—he always likes to help. Not once has he been rude to me. Not once has he used harsh words. And I am always the one wanting to be the master, while he is only the servant and errand boy." Regi no longer paid attention to the pain in her feet... (Kijne, 1958, p. 29)

This excerpt shows guilt emerging as deep introspection within Regina. Her remorse is not just emotional but marks significant moral and social awakening. For the first time, Regina views herself critically, recognizing the injustice of her attitude toward Tomi. Her guilt is conveyed through sobbing, losing composure, and ignoring physical pain. These reactions show she has stepped outside the superiority that justified her commanding and reprimanding. This shift is prompted by recalling Tomi's kindness: he is consistently helpful, gentle, and never harsh. Such reflections compel her to reevaluate her behavior, admitting she treated him unfairly and abusively.

The statement, "I always wanted to be the master, while he was only the servant and errand boy," reveals Regina's awareness of the social hierarchy she perpetuated. Her remorse signifies the collapse of her internalized colonial worldview—one that normalizes dominance and subordination. In Goldmann's genetic structuralism, this moment signifies a rupture in her consciousness and the emergence of a possible consciousness—an alternative, more egalitarian and humane way of understanding her relationship with Tomi.

Regina's guilt illustrates that ideologies of domination not only oppress the marginalized but distort the emotional and relational life of those in dominant positions. Her newfound awareness marks the beginning of a shift in power dynamics between herself and Tomi, paving the way for reconciliation. Thus, Regina's remorse is not just part of her emotional journey; it functions as a narrative mechanism reinforcing the central themes of the story—friendship, equality, and restoration of human relationships damaged by social hierarchies. Her guilt becomes the moral foundation for seeking Tomi, trying to mend the relationship, and ultimately rebuilding a bond rooted in affection and mutual respect.

(c). *Expressions of Grief*

The following passage illustrates the profound sorrow experienced by Regina after losing Tomi and confronting an overwhelming sense of emotional emptiness:

The flower garden was silent. The sky was merely gray, covered with clouds. Regina wandered alone in search of something to play with. But she could not play. Everything felt difficult. The flowers no longer sparkled. Even Pit the cassowary walked about pecking gloomily. Regi lamented, "Today is not good at all. I don't enjoy playing ball. I don't want to pick flowers. I am tired of running around with Pit. Never have I felt such great sorrow. Where is that beautiful light now? Where is the golden city?" (Kijne, 1958, p. 15)

This passage portrays grief as both an emotional response and a symbolic state reshaping Regina's perception of the world. Her sense of loss renders her surroundings colorless, lifeless, and joyless, suggesting grief as an existential experience disrupting her worldview. Descriptions of nature as "gray," "silent," and "no longer sparkling" reflect Regina's emotional landscape. In Goldmann's terms, shifts in environmental tone within a text often mirror a character's psychological structure. Here, Regina's grief extinguishes the vibrancy and beauty she once perceived. Even animals, typically lively companions, appear burdened, reinforcing how her sorrow reframes everything.

Regina's repeated complaints—her inability to enjoy cherished activities—underscore the absence of an emotional anchor: Tomi. Her grief emerges from realizing she lost a meaningful, previously taken-for-granted relationship. Thus, sorrow marks the second stage of her moral transformation, following anger and preceding reconciliation. Her question, "Where is the golden city?" reveals grief extends beyond personal loss, signifying the collapse of an ideal image of happiness. The Golden City, symbolizing peace, friendship, and balance, feels distant and unattainable, indicating Regina's understanding that such an ideal world only emerges through nurturing relationships—particularly with Tomi.

In the broader ideological context, this scene reinforces Kijne's worldview that emotional and social barrenness arises from ruptured relations of compassion. Regina's grief critiques the unequal relational structure she upheld and paves the way for new consciousness: realizing a beautiful life is built not on dominance but on friendship, cooperation, and mutual care.

(d). *Expressions of Joy*

The following passage captures the climactic moment of joy experienced by Regina and Tomi as they finally enter the Golden City together:

"Is it still far?" asked Tomi. "No, just a little more. Once we turn that corner you will see the city. Look! Its light is already shining on the rocks beyond that ledge!" "Oh Tomi, thank goodness, the gate is open." Hand in hand, the two of them ran toward the golden gate. The guard laughed. "Ha, so you've found Tomi? Good. Rejoice together, yes! Go on in, don't be afraid." The most beautiful golden light shone from inside the gate. The air was filled with melodious sounds. The two children just stood there, letting the joy seep deeper into their hearts. Tom sighed, "Do you dare go in?" Regina whispered, "Do you?" Then, hand in hand, they stepped under the archway and stood inside the Golden City. (Kijne, 1958, p. 44)

This passage portrays happiness not as fleeting delight but as a profound emotional and symbolic experience—restoring their relationship, fulfilling their quest, and embodying the author's worldview. The joy felt by Regina and Tomi marks the culmination of their journey and reconciliation after conflict.

First, this joy emerges from renewing their friendship. Their clasped hands, running together, and mutual question—"Do you dare go in?"—signal a transformed relationship grounded in equality and support. The master-servant hierarchy has dissolved; what remains is partnership and loyalty. In Goldmann's terms, this scene illustrates the realization of a possible consciousness—an alternative social relation that is more just, humane, and grounded in solidarity.

Second, the depiction of the Golden City—radiant light, melodious air, and overwhelming beauty—shows that happiness in this narrative is symbolic. The city is not just an imagined place but represents the author's highest values: peace, companionship, and harmonious living. Joy arises only when the children enter this symbolic space together, emphasizing that unity and affection are prerequisites for an ideal world.

Third, their happiness carries a dimension of moral transcendence. After passing through anger, remorse, and sorrow, the children experience inner exaltation marking their ethical transformation. The golden light that "seeps into their hearts" symbolizes moral illumination: true human connection is built through forgiveness and friendship.

Moreover, the brief dialogue—"Do you dare go in?"—demonstrates that happiness is not simply bestowed; it is a moral decision. Their choice to enter together suggests that an ideal society can be achieved through cooperation and courage. From the perspective of genetic structuralism, this symbolic act reflects the homology between narrative structure and the social vision the author upholds: a world where love transcends race, status, and historical domination.

Thus, the happiness depicted in this scene represents the apex of the story's emotional and social dialectic. It affirms that friendship and humanity are foundational principles of the Golden City—both as an imagined destination and as an ethical vision for Papua's future.

B. *Social Human Facts*

(a). *Social Activities*

The following passage illustrates how Regina and Tomi engage in a social activity involving cooperation as they climb a rocky terrain:

The path upward was still just as steep. The rocks were still large and sharp. Regina recognized all the thorns. But now the two of them helped each other. Regina supported Tom as he climbed the high rocks. Then Tom reached out his hand to pull Regina up. Tom called out, "Watch out, there are many thorns!" Regina shouted, "Be careful, that rock is about to fall...." (Kijne, 1958, p. 43)

This excerpt depicts a social activity that reflects values of togetherness, solidarity, and cooperation between Regina and Tomi. Their joint effort in climbing the rocky slope shows their relationship is now based on mutual dependence and

shared purpose, not hierarchy. The scene illustrates how social actions like helping each other reshape interpersonal relations.

This mutual assistance marks a significant shift in their relationship dynamics. Previously, Regina was dominant and Tomi submissive; now, both support each other physically and emotionally. Regina helps Tom climb, and Tom pulls Regina upward. Cooperation thus dismantles their earlier power asymmetry. The activity reflects what Goldmann describes as a collective subject, not limited to large groups but also small interactions expressing shared values and goals. Regina and Tomi form a small collective subject aiming for safety and companionship, signaling a new collective consciousness beyond class or ethnic distinctions. Brief exchanges like “Watch out, there are many thorns!” and “Be careful, that rock is about to fall!” highlight communication as integral to social activity. Their dialogue is now expressions of reciprocal concern, demonstrating healthy relationships built on mutual care, not domination. This scene serves as a microcosm of the author’s ideological vision. For Kijne, cooperation and mutual assistance are foundational values for building the “Golden City”—a symbol of an ideal society grounded in peace and justice.

Thus, this activity reflects the author’s aspirational worldview for Papua: a society where cultural and social differences are transcended through solidarity. The data show that social activities in the story function not merely as narrative elements but as expressions of the ideal social structure envisioned by the author. Regina and Tomi’s cooperation exemplifies a worldview placing love, equality, and communal harmony at the core of human life.

(b). Cultural Activities

The following passage depicts Regina’s interaction with *the kris ekor kipas (the fan-tailed kris)*, a Papuan fan-tailed bird that offers her a sacred feather and a magical incantation to guide her toward the Golden City:

Where is that beautiful light now? Where is the golden city? Suddenly *the fan-tailed kris* leapt onto the wooden fence—right in front of Regina’s nose! It whispered to her: “You can reach it, you can reach that city. Truly! I want to help you. Listen carefully! Just pluck one feather from my tail. Be careful, don’t lose it. If you ever need anything, just say: ‘Wirewit,’ while holding the feather.” Regina felt as though she were dreaming... She plucked the feather... and the bird had vanished. (Kijne, 1958, p. 41)

This passage reveals a strong cultural dimension, marked by Papuan folklore, natural symbolism, and magical practices serving narrative and ideological functions. Regina’s encounter with the fan-tailed kris is not just fantastical but represents living Papuan cultural traditions expressed through animals, symbols, and rituals. The fan-tailed kris—a bird native to Papua—is a crucial cultural signifier. It acts as a spiritual mediator, guiding Regina. In Papuan traditions, animals possess symbolic power, often serving as guardians, omens, or messengers. This scene reflects Papuan communities’ cultural imagination, conceptualizing humans’ relationship with nature as harmonious and sacred. The offering of a feather as a talisman and the mantra “Wirewit” mirror local practices. Across Indonesia, natural objects like feathers, stones, or leaves are imbued with transformative qualities.

The feather marks Regina’s entry into Papua’s cultural world, suggesting reaching the Golden City—an ideal social order—requires understanding and accepting local values. Her transformation is moral and cultural. This cultural activity highlights local wisdom as a guide to a better life. Regina cannot reach the Golden City alone; she needs help from a symbolic being representing indigenous knowledge. In Goldmann’s terms, this act reveals collective consciousness: true happiness is attainable only through harmonious relations between humans and nature, and newcomers and Indigenous inhabitants.

The passage affirms the author’s view on respecting local wisdom. As a missionary and educator, Kijne understood Papuan society couldn’t be shaped solely by Western values; it required synergy with long-standing traditions. This scene conveys an ideological message that modernity must coexist with tradition. Thus, this cultural activity reflects Papuan society’s social and symbolic structures while showing how local culture becomes the foundation for characters’ spiritual and moral journey. The integration of folklore, natural symbols, and ritual language underscores that the Golden City—Kijne’s envisioned ideal—can only be reached through appreciating and preserving local cultural heritage.

C. Collective Subject

“Regi!” Her mouth fell open, her eyes widened, and for a few moments she stood completely frozen. Then she turned around slowly, sobbing uncontrollably. “Oh, it is my fault, I am a bad child. I was so angry at Tomi.” As Regi descended the mountain, many thoughts arose in her heart. “Good Tom. He always likes to help. Not once has he been rude to me. Not once has he used harsh words. And I am the one who always wants to be the master and the lady, while he is only the servant and the helper”. (Kijne, 1958, p. 29)

This passage illustrates the collective subject within Kota Emas’s narrative. The “master and lady” versus “servant and helper” dichotomy is not just Regina’s personal view; it reflects the class hierarchy and power dynamics in colonial Papua. Regina’s self-awareness as “the master and the lady” signifies her position in the dominant European settler group, historically over Indigenous Papuans. Tomi represents the subaltern class, relegated to helper or servant roles.

This dynamic is not just interpersonal but arises from the collective consciousness of larger social groups, as Goldmann describes the collective subject: an ideological structure shared by a social group, manifesting in characters’ actions and attitudes. The passage shows the internalization of colonial ideology. Regina assumes command entitlement, while Tomi’s role is obedience, a product of her social structure, not individual reasoning.

Thus, Regina represents the colonial group, framing Indigenous people as subordinates. Regina's introspection—realizing Tomi's kindness and helpfulness—signals her inherited collective consciousness breaking down and a more egalitarian awareness emerging. In Goldmann's framework, this reflects a move toward possible consciousness, transcending class-based worldviews. Regina's transformation shows ideology is not static. Their hierarchical relationship shifts toward a horizontal, friendship-based dynamic, critiquing the colonial social order. Kijne suggests equitable relations emerge when dominant group members recognize and reshape old structure injustices. Thus, this passage represents Goldmann's collective subject concept: characters act with social structures, ideologies, and collective consciousness of their groups. The move from "master-servant" relations to genuine friendship marks the social structure transformation idealized by the author.

D. Worldview

"...Indeed, if you are grieving, you will surely find joy in the city. But... I have heard from Kris the fan-tailed bird that you have a friend. Why have you come here alone? Why did you not call your companion first? He must accompany you, must he not? 'To grieve together, to rejoice together.' If you were to enter now, you would not be able to find happiness". (Kijne, 1958, p. 29)

This passage reveals the worldview the author constructs in Kota Emas, centered on friendship, togetherness, and solidarity as foundations of human happiness. The phrase "To grieve together, to rejoice together" encapsulates the author's stance on forming human relations.

First, the gatekeeper's statement underscores that happiness cannot be attained individually. Regina wishes to enter the city to soothe her sorrow, yet the gatekeeper prevents her because she arrives alone. In Goldmann's framework, this reflects that moral and social resolutions cannot be achieved in isolation; individuals exist within relational networks. Kota Emas, as a symbol of the ideal world, is accessible only to those in healthy, egalitarian relationships.

Second, the text emphasizes friendship as a moral prerequisite for genuine happiness, contrasting with the colonial consciousness earlier portrayed—where Regina was "master" and Tomi the "servant." By requiring Regina to enter the city only with Tomi, the narrative rejects hierarchical social structures as incompatible with the ideal world. Friendship signals ethical equality forming the basis of Kota Emas.

Third, the assertion that she "would not be able to find happiness" if she entered alone suggests happiness is not determined by external circumstances but by one's internal state and social relations. From Goldmann's perspective, this aligns with the notion that literary works express a collective consciousness concerning values necessary for an ideal life.

Fourth, the symbolic figures—the gatekeeper, Kris the fan-tailed bird, and Kota Emas itself—demonstrate how the author weaves local cultural motifs with universal moral principles. Kota Emas is not merely fantastical but represents the ethical ideal: happiness must be cultivated within the community rather than individualistic pursuit.

Thus, this passage illustrates that the author's worldview centers on moral messages of humanity, equality, and solidarity. In the story, Kota Emas functions not only as an imaginative space but also as a symbolic vision of an ideal society—one attainable only through respectful, compassionate, and mutually supportive social relationships.

E. Narrative Structure

(a). Cultural Opposition: Imagination vs. Reality

"Oh Tom, look over there!" whispered Regi to Tom. "Across that place. Behind the forest! It looks like a city of gold!" The faces of the two children gleamed in the golden light radiating from the horizon. Houses and walls, gates and towers appeared—all golden, shining brilliantly. "Regi..." "Yes, Tom!" "Do you know what I wish for? I want to have wings so that I can fly. I want to fly there, to that bright city. Beyond the distant hills, near the sun! I think the Good Lord lives there. Ah, if only I could enter it, even for just one hour!" (Kijne, 1958, p. 7)

This passage illustrates imagination's crucial role in the story's structure. Kota Emas appears as a utopian vision, existing only in the children's imagination, contrasting the real world of conflict with an ideal world of purity and joy. This vision emphasizes the tension between reality and aspiration. Regina and Tomi live in an unequal society but imagine a city "gleaming in golden light," symbolizing hope.

This contrast highlights the difference between their reality and the ideal world they desire. The city, described as "brilliant" and "near the sun," is a transcendental space in the children's consciousness, aligning with Goldmann's concept of a possible world, a vision of a just, humane existence. Tomi's wish to "have wings" to fly there shows the ideal world as initially unattainable, accessible only through fantasy. Imagination motivates the characters toward moral transformation. The notion that "the Good Lord lives there" reinforces Kota Emas as a moral ideal, embodying divine values.

Thus, the children's vision represents an ideological construct of life's ultimate purpose. Imagination vs. reality is a cultural opposition in the story, with imagination offering hope and vision, while reality presents conflict. The narrative moves from imagination to actualization, making this opposition central to the plot. In sum, imagination is integral to the story's ideology, suggesting an ideal life can be envisioned and realized through moral transformation and restored relationships.

(b). Narrative Structure – The Hero Figure

One morning Regina said, "Let's go to the forest, come on!" "But you can't, Regi. Your parents have forbidden it, haven't they?" "Not far, just a little bit. And you are always allowed to go." The two of them left the flower garden and entered the forest... "Oh no! Look over there! A big boar! Across the river." Tom, the good-hearted boy, turned back—astonishingly brave. He, too, was afraid of the boar, but he refused to abandon his friend. Quickly he helped Regi to her feet. Holding hands, the two friends ran back toward the garden. Pit the cassowary did not understand why his two friends were sitting pale-faced on the grass. (Kijne, 1958, p. 9)

This passage presents Tomi as a heroic figure through courage and self-sacrifice, revealing his capacity to overcome fear to protect Regina, reflecting the moral values promoted by the author. Tomi's bravery is clear when he faces the wild boar despite fear. The line "he was also afraid of the boar, but he refused to abandon his friend" shows his heroic action stems from responsibility and concern for others. Within the story's moral framework, courage is acting despite fear for another's sake. Tomi emerges as both guardian and rescuer, warning Regina, helping her up, and guiding her to safety, contrasting earlier depictions of him as a "jongos" or servant.

This marks a crucial shift, showing heroic qualities often emerge from those seen as inferior within colonial structures. Within genetic structuralism, Tomi's act reflects a shift in consciousness. He is no longer a passive social subject but a moral agent directing the narrative's ethical trajectory, moving from a subaltern role to a moral protagonist, establishing equality with Regina. The scene critiques colonial hierarchies, emphasizing courage and moral worth are not determined by race, skin color, or class.

This scene also sets the stage for Regina's transformation. Tomi's bravery compels her to recognize her dependence on him, leading to reflection, remorse, and reconciliation in later sections. Thus, Tomi's act is not just physical intervention but a catalyst for emotional and moral development throughout the story. The event reinforces a central value of *Kota Emas*: friendship and mutual protection. Tomi's heroism shows the Golden City is reached not through adventure alone, but moral courage and compassion. Tomi's heroism embodies the author's worldview that humanity, bravery, and solidarity transcend social boundaries constructed by colonial society. Tomi's courage symbolizes the potential of marginalized individuals to become agents of change within the narrative's social structure and emotional progression.

V. CONCLUSION

The analysis of *Kota Emas* by Izak S. Kijne through genetic structuralism shows the story's deep link between characters' experiences and the colonial social structure of Papua. The human facts, collective subject, and cultural oppositions in the narrative highlight social tensions between Western dominance and the subaltern Papuan communities, while opening a space for a shift toward egalitarian relationships. Kijne's worldview emphasizes friendship, solidarity, and compassion as moral foundations to overcome class, race, and cultural differences. Moreover, the story's structure—symbolized by the Golden City and conflicts and reconciliations—creates a homology between the text and Papua's socio-historical reality. Through understanding and explanation, the narrative affirms the author's vision of harmony, cooperation, and humanity as essential for a just, peaceful society. These findings show *Kota Emas* is not just a children's tale but an ideological text reflecting Papuan society's aspirations and contributing significantly to Eastern Indonesian literary studies.

REFERENCES

- [1] Afiah, N., Arafah, B., & Abbas, H. (2022). Burmese Women Portrait under the British Imperialism in Orwell's *Burmese Days*. *Journal of Language Teaching and Research*, 13(1), 213–219.
- [2] Arafah, B. (2018). Incorporating the Use of Literature as an Innovate Technique for Teaching English. *The 1st Annual International Conference on Language and Literature, KnE Social Sciences*, 2018, 24–36.
- [3] Arafah, B., Asri., D., Hanafiah, W., Ruslin, Ilmi, N., Pratiwi, A., & Astiantih, S. (2024). Female Existentialism in *Monkey Grip* by Helen Garner. *Journal of Language Teaching and Research*, 15(4), 1238–1246.
- [4] Arafah, B., Hasyim, M., Aman, R. bin, Agussalim, A., & Latief, M. R. A. (2026). Language as a Gateway to Authentic Tourism: Leveraging Bugis-Makassar Expressions in South Sulawesi's Tourism Industry. *Theory and Practice in Language Studies*, 16(2), 597–604.
- [5] Arafah, B., Hasyim, M., Khaerana, A. S. A., Soraya, A. I., Ramadhani, R., Ismail, A., & Takwa. (2023a). The Digital Culture Literacy of Generation Z Netizens as Readers, Producers and Publishers of Text on Social Media. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3), 112–123.
- [6] Arafah, B. & Hasyim, M. (2023a). Digital Literacy: The Right Solution to Overcome the Various Problems of Meaning and Communication on Social Media. *Studies in Media and Communication*, 11(4), 19–30.
- [7] Arafah, B. & Hasyim, M. (2023b). Digital Literacy on Current Issues in Social Media: Social Media as a Source of Information. *Journal of Theoretical and Applied Information Technology*, 101(10), 3943–3951.
- [8] Arafah, B., & Kaharuddin. (2019). The Representation of Complaints in English and Indonesian Discourses. *Opción*, 35(24), 501–517.
- [9] Arafah, B., Kaharuddin, Hasjim, M., Arafah, A. N. B., Takwa, & Karimuddin. (2023b). Cultural Relations among Speakers of South Halmahera Languages. *Theory and Practice in Language Studies*, 13(1), 168–174.
- [10] Arafah, B., Rofikah, U., Kaharuddin, A., Room, F., Takwa, & Purwarno, P. (2023c). Evaluating the Suitability of Printed Materials in the Context of Online Learning. *XLinguae*, 16(2), 16–32.
- [11] Arafah, B., Room, F., Suryadi, R., B., L. O. M. I. H., Juniardi, Y., & Takwa. (2023d). Character Education Values in Pullman's *The Golden Compass*. *Journal of Language Teaching and Research*, 15(1), 246–254.

- [12] Arafah, B., Sirajuddin, J., Thayyib, M., Room, F., Takwa, & Anayati, W. (2023e). Emotional Management of Defoe's *Robinson Crusoe's* Main Character. *Journal of Language Teaching and Research*, 14(5), 1414-1421.
- [13] Arafah, B., Hasyim, M., Arafah, A. N. B., Taqdir, Nursidah, Ruslin, & Arafah, F. R. B. (2025). Local Wisdom and Leadership Values in Bugis Governance: Insights into Indigenous Political Ethics in South Sulawesi. *Scientific Culture*, 11(3.2), 224-234.
- [14] Arifin, M. B., Arafah, B., & Kuncara, S. D. (2022). Dayak's Sociocultural Situation through Locality in Lumholtz's *Through Central Borneo* Travel Writing. *Theory and Practice in Language Studies*, 12(12), 2695-2703.
- [15] Arifuddin, A., Arafah, B., Abbas, H., Lestari, W., JUSDalyana, Tang, M. A., & Malik, A. N. M. (2024). Racial Injustice against Blacks in American Society as Represented in Wright's *Native Son*. *Theory and Practice in Language Studies*, 14(12), 3938-3946.
- [16] Arnawa, I. G. N. E. V., & Arafah, B. (2023). Students' Self-Regulated Strategies in Approaching Second Language Writing. *Theory and Practice in Language Studies*, 13(3), 690-696.
- [17] Asri, D., Arafah, B., Sahib, H., & Abbas, H. (2023). Male Domination in Helen Garner's *Monkey Grip*. *Theory and Practice in Language Studies*, 13(7), 1651-1658.
- [18] Asriyanti, S., Arafah, B., & Abbas, H. (2022). The Representation of Women's Dependence on Men in *Little Women*. *Theory and Practice in Language Studies*, 12(4), 790-796.
- [19] Baa, S., Wardani, S. B., Iskandar, Weda, S., & Arafah, B. (2023). Lexical Metaphors in Westlife's Selected Song Lyrics. *XLinguae*, 16(1), 132-154.
- [20] Effendy, C., Kurniawan, S., Arafah, B., & Hidayat, S. (2025). Ghosts in Pontianak Malay Beliefs: Origins, Teachings, and Evolving Functions. *Scientific Culture*, 11(3.1), 1476-1492.
- [21] Fadillah, N., Arafah, B., & Abbas, H. (2022). The Act of Slavery in the 20th Century as Reflected in Samuel Beckett's *Waiting for Godot*. *Journal of Language Teaching and Research*, 13(1), 66-72.
- [22] Goldmann, L. (1977). *Essays on Method in the Sociology of Literature*. London: Tavistock.
- [23] Halil, N. I., Arafah, B., Saputra, I. G. P. E., Hasyim, R. S., Sarmadan, Takwa, & Karma, R. (2024). Preservation of Tolaki Mekongga Language Through Merdeka Curriculum-Based Local Subject Teaching Modules. *Journal of Language Teaching and Research*, 15(3), 960-971.
- [24] Hasjim, M., Arimi, S., Kushartanti, B., Arafah, B., & Taqdir. (2026). Language Transformation in Social Media and Its Impact on the Linguistic Identity of Indonesian Youth. *Journal of Language Teaching and Research*, 17(1), 296-305.
- [25] Hasyim, M. & Arafah, B. (2023a). Semiotic Multimodality Communication in the Age of New Media. *Studies in Media and Communication*, 11(1), 96-103.
- [26] Hasyim M. & Arafah, B. (2023b). Social Media Text Meaning: Cultural Information Consumption. *Wseas Transactions on Information Science and Applications*, 20(2023), 220-227.
- [27] Idhan, M., Jabir, Muh., Ruslin, R., Asmawi, M. N., Arafah, B., Ubadah, Ubay, & Fatima. (2026). Teachers' Experiences With Mobile Apps for Arabic Language Learning: Insights From Islamic Schools in Palu, Indonesia. *Journal of Language Teaching and Research*, 17(1), 338-350.
- [28] Iksora, Arafah, B., Syafruddin, S., Muchtar, J., & Lestari, P.A. (2022). Typos' Effects on Web-Based Programming Code Output: A Computational Linguistics Study. *Theory and Practice in Language Studies*, 12(11), 2460-2469.
- [29] Jaelani, A., Arafah, B., Abbas, H., & Yudith, M. (2024). Environmental Themes in Michael Punke's *The Revenant*: An Ecocritical Analysis of Human-Nature Interaction. *Theory & Practice in Language Studies*, 14(10), 3312-3319.
- [30] Jamiluddin, Suparti, Arafah, B., & Marhum, M. (2026). Language, Power, and Symbolic Violence in Classroom Communication: Insights From Indonesian Culture. *Journal of Language Teaching and Research*, 17(2), 735-746.
- [31] JUSDalyana, Arafah, B., Abbas, H., Lestari, W., Arifuddin, A., Tang, M. A., & Malik, A. N. M. (2024). Representation of Patriarchal Ideology and Its Negative Effect on the Characters in Naomi Alderman's *The Power*. *Theory and Practice in Language Studies*, 14(9), 2880-2888.
- [32] Kaharuddin, Ahmad, D., Mardiana, Latif, I., Arafah, B., & Suryadi, R. (2024). Defining the Role of Artificial Intelligence in Improving English Writing Skills Among Indonesian Students. *Journal of Language Teaching and Research*, 15(2), 568-578.
- [33] Kaharuddin, Arafah, B., Nurchahmi, S., Sukmawaty, Rahman, I. F., & Juniardi, Y. (2023). Exploring How Reading Aloud and Vocabulary Enrichment Shape English Speaking Skills Among Indonesian Learners of English. *World Journal of English Language*, 13(8), 436-445.
- [34] Kijne, I. S. (1958). *Kota Emas (Golden City)* (Fourth edition). Gronigen: J. B. Wolters.
- [35] Kuswanti, W. H., Arafah, B., Budiman, A. N. A., Ali, T., Fatsah, H., & Room, F. (2023). Students' Perception of Explicit and Implicit Methods in Learning Tenses in SMP DDI Mangkoso. *Theory and Practice in Language Studies*, 13(6), 1473-1482.
- [36] Lestari, W., Arafah, B., Abbas, H., JUSDalyana, Arifuddin, A., Tang, M. A., & Malik, A. N. M. (2025). Gender Representation in Baker's *Princess Between Worlds*. *Theory and Practice in Language Studies*, 15(1), 55-63.
- [37] Manugeran, M., Arafah, B., Purwarno, P., Siwi, P., Ekalestari, S., & Wulan, S. (2023). An Ecoliterature Approach to Environmental Conservation: Take Four Selected Literary Works as Examples. *Theory and Practice in Language Studies*, 13(5), 1318-1327.
- [38] Mare, B. S., Arafah, B., Abbas, H., Rano, W. A., Almahdali, U. H., Wicaksono, C. K., Sari, R. A., & Malik, N. F. (2025). The Influence of Social Norms and Etiquette on the Behavior and Decision-Making of Characters in Jane Austen's *Pride and Prejudice*. *Theory and Practice in Language Studies*, 15(12), 4006-4016.
- [39] Mokoginta, K. & Arafah, B. (2022). Negotiation in Indonesian Culture: A Cultural Linguistic Analysis of Bahasa Indonesia Textbooks. *Theory and Practice in Language Studies*, 12(4), 691-701.
- [40] Mutmainnah, Arafah, B., & Pattu, A. (2022). Racial Discrimination Experienced by Black People as Reflected in Langston Hughes's Poems. *Journal of Language Teaching and Research*, 13(2), 350-356.
- [41] Pratiwi, D. S., Sarwono, S., & Lubis, B. (2017). Analisis Novel *Perahu Kertas* Karya Dewi Lestari: Kajian Strukturalisme Genetik [Analysis of Dewi Lestari's Novel *Perahu Kertas*: A Genetic Structuralism Study]. *Jurnal Ilmiah KORPUS*, 1(1), 32-38.

- [42] Purwaningsih, Y. R., Arafah, B., Abbas, H., & Arafah, A. N. B. (2020). Psychological Disorder of Emma Bovary as a Wife toward Sexuality Dissatisfaction in Gustave Flaubert *Madame Bovary*. *Talent Development & Excellence*, 12(3s), 1382-1389.
- [43] Radjaban, R. Y., Humanika, E. S., & Arafah, B. (2025). A Holistic Study of Merap Language Mapping in Relation to the Term Punan-Merap. *Journal of Language Teaching and Research*, 16(3), 957-965.
- [44] Rano, W. A., Arafah, B., Abbas, H., Almahdali, U. H., Wicaksono, C. K., Sari, R. A., Malik, N. F., & Mare, B. S. (2026). Social Class and Stratification in the Edwardian Era as Portrayed in Forster's *A Room With A View*. *Theory and Practice in Language Studies*, 16(2), 639-648.
- [45] Rukiah, S. (2019). Nilai Tanggung Jawab dalam Novel Laskar Pelangi Karya Andrea Hirata: Tinjauan Strukturalisme Genetik [The Value of Responsibility in the Novel Laskar Pelangi by Andrea Hirata: A Review of Genetic Structuralism]. *Ksatria Jurnal Kajian Bahasa dan Sastra*, 1(1), 43-56.
- [46] Siwi, P., Arafah, B., Wulan, S., Purwarno, P., Ekalestari, S., & Arafah, A. N. B. (2022). Treatment of Nature: An Ecocriticism Approach in 'Komat Kamit' of Tejo and Kamba's *Tuhan Maha Asik*. *Theory and Practice in Language Studies*, 12(7), 1278-1285.
- [47] Suhadi, J., Arafah, B., Makatita, F. P., Abbas, H., & Arafah, A. N. B. (2022). Science and Society: The Impact of Science Abuse on Social Life in Well's *The Invisible Man*. *Theory and Practice in Language Studies*, 12(6), 1214-1219.
- [48] Sunyoto, F. G., Arafah, B., Yudith, M., Mokodompit, G. P., & Asnawi, A. E. F. (2022). The Native American Father's Parenting Style in John Ernst Steinbeck's *The Pearl*. *Theory and Practice in Language Studies*, 12(12), 2551-2558.
- [49] Takwa, Arafah, B., Hasyim, M., Akhmar, A. M., & Taqdir. (2025). Cultural Value in Mowindahako Ritual Speeches of Tolaki Mekongga Ethnic Group in Kolaka, Indonesia. *Theory and Practice in Language Studies*, 15(6), 1857-1866.
- [50] Takwa, Arafah, B., Hasyim, M., Akhmar, A. M., & Taqdir. (2026). Linguistic Units and Cultural Meanings in the Ritual Speech of Mowindahako Among the Tolaki Mekongga Ethnic Group. *Journal of Language Teaching and Research*, 17(2), 665-676.
- [51] Takwa, Arafah, B., Hasyim, M., Akhmar, A. M., & Taqfiah, S. J. A. (2024a). Symbolic Meanings in the Mowindahako Ritual of the Tolaki Mekongga Kolaka Ethnic. *International Conference on Medical Imaging, Electronic Imaging, Information Technologies, and Sensors (MIEITS 2024)*, 13188(131880W), 152-156.
- [52] Takwa., Arafah, B., Hasyim, M., & Akhmar, A. M. (2024b). Cultural Imagery of Tolaki Mekongga Community of Kolaka in Mowindahako Ritual. *Theory and Practice in Language Studies*, 14(3), 763-770.
- [53] Takwa, T., Arafah, B., Taqdir, T., Sarmadan, S., Tundreng, S., Pageyasa, W., Hariyadi, S., Rasid, R., Inpresani, W. O., & Mofu, H. (2026). Local Knowledge for Coastal Resilience: Adaptive Strategies of Bajo Fishers in Pomalaa, Indonesia. *Scientific Culture*, 12(1.1), 1336-1347.
- [54] Taqdir, Arafah, B., & Nursidah. (2026). Perceptions of Cultural Representation in Japanese Language Textbooks: Insights From Students and Teachers in Indonesian Higher Education. *Journal of Language Teaching and Research*, 17(2), 622-633.
- [55] Taqdir, Arafah, B., Nursidah, & Afisha, A. W. O. N. (2025). Gender Representation in Foreign Language Textbooks: Trends, Gaps, and Implications for Multilingual Teaching. *XLinguae*, 18(4), 225-242.
- [56] Taqdir. (2025). Cultural Stereotypes in Foreign Language Textbooks: A Systematic Review of Visual Representation. *Theory and Practice in Language Studies*, 15(6), 1949-1958.
- [57] Usman, S., Arafah, B., Marhum, M., Munir, S., Budi, & Tadeko, N. (2024). The Effect of Technological Pedagogical Content Knowledge (TPACK) on Rural Area Students' English Writing. *Journal of Language Teaching and Research*, 15(3), 884-892.
- [58] Yudith, M., Arafah, B., Sunyoto, F. G., Fitriani, Rostan, R. B., & Nurdin, F. E. (2023). The Representation of Animalism Issue in Sewell's *Black Beauty*. *Theory and Practice in Language Studies*, 13(1), 108-116.
- [59] Yudith, M., Arafah, B., Abbas, H., Jaelani, A., Sunyoto, F. G., & Arafah, A. N. B. (2024). Social Issues during the Victorian Era Lead to the Formation of Nice Girl Syndrome in Gaskell's *Ruth*. *Journal of Language Teaching and Research*, 15(3), 822-832.



Insum Malawat obtained her Doctoral degree in Indonesian Language Education at the State University of Malang (UM), Indonesia 2017. Her Master's degree in Indonesian Literature was obtained at Gadjah Mada University (UGM) Indonesia in 2010. Her Bachelor's Degree in Indonesian Language Education was obtained at Cenderawasih University (Uncen).



Hengki Mofu is a Dean at the Faculty of Teacher Training and Education, University of Papua. He teaches several linguistics courses, such as phonology, morphology, syntax, semantics/pragmatics, and psycholinguistics. He has also researched Rising and Falling intonation in yes/no questions among English Study program students and Morphological Process through Inflectional Suffixation in English and Muna Language: A Contrastive Study.



Burhanuddin Arafah obtained his PhD in English (Australian) literature at the University of Newcastle, Australia, in 2003. He earned his Master's degree in American literature at Gadjah Mada University (UGM) Yogyakarta of Indonesia in 1995 and his Bachelor's degree in English literature at Hasanuddin University (UNHAS) in 1988. He is currently a full professor of English literature at the English Literature Study Program, Faculty of Cultural Sciences of UNHAS, Indonesia.

Symbolic Articulations of Post-Traumatic Memory in Women's Poetry: Fire, Ashes, and Self-Transformation in Atwood and Al-Malaika

Fatimah A. Alwadaani

Department of English Language, College of Arts, King Faisal University, Al Ahsa, Saudi Arabia

Abstract—This paper explores poetry's capacity to transform post-traumatic memory into symbolic articulation in two major collections of prominent poets, Margaret Atwood and Nazik Al-Malaika. Both poets come from different cultures yet explore similar emotional experiences. Drawing on the psychological approach and trauma theories of Freud and Jung, this paper probes the psychological concepts of inner entrapment, fragmentation, transformation, and individuation. While prior studies have explored trauma in women's poetry, limited attention has been paid to the symbolic articulation of post-traumatic memory across different cultures. By focusing on language and imagery in the selected works, the study aims to demonstrate how post-traumatic memory can distort the perception of both inner and external realities. In addition, it argues that the employment of symbols such as fire and ashes reflects the poets' emotional bonds and psychological development throughout the poignant articulation of traumatic memories. The findings reveal a significant resemblance between the two poets in the symbolic expression of traumatic memory despite differences in time and cultural contexts. Both Atwood and Al-Malaika demonstrate psychological maturity through true self-acceptance and resilience. This paper contributes to studies of trauma and memory by enhancing the understanding of women's poetic engagements with private traumatic memory as a transformative symbolic practice. Future research examining additional linguistic techniques in the poetry of diverse poets could enrich interdisciplinary scholarship and further illuminate poetry's role in post-traumatic self-reconstruction.

Index Terms—comparative literature, poetic symbol, post-traumatic memory, psychological reading

I. INTRODUCTION

Post-traumatic memory is a profound psychological burden that emotionally disrupts self-representations and complicates straightforward narratives of lived experiences. According to Nadal and Calvo (2014), "It is not the first act which is traumatic, it is the internal reviviscence of the memory that becomes traumatic" (p. 3). Emerging from psychological studies and complex experiences of humans suffering, trauma theory emphasizes the fragmented and resistant nature of traumatic memory. The trauma theorists argue that these memories often defy linear narration, which makes literary forms -especially poetry- a suitable medium for articulation. Poetry provides a space for repressed experiences and memories through its reliance on symbolism and non-linear structure. Ulrich Baer (as cited in Armstrong, 2020, p. 543) notes, "enacting rather than directly communicating its truth, poetry registers the type of experience that an individual, after suffering a trauma, may be compelled to reenact but cannot recall at will". This insight validates the use of symbolic imagery in Atwood's and Al-Malaika's poetry in order to articulate post-traumatic memory.

Painful memories make it difficult for writers to emotionally explain themselves in their literature. Therefore, they rely on figurative language and representations to demonstrate their experiences and reveal obscured areas of their psychological situation. This strategy can be seen in the poetry of Margaret Atwood and Nazik Al-Malaika, who employ the symbols of burning and ashes to express the self-fragmentation that occurs due to a post-traumatic memory. The emotional experiences that the two poets went through left deep scars in their psyches and vividly appear through their imagination. Therefore, psychological theories, mainly trauma theory, were used to explain what lies beneath and works as a catalyst for their poetic expressions.

Although better known for her novels, Atwood is also one of Canada's premier poetic voices, expressing themes of alienation, myth, power, women, and survival with an intellectual sharpness and emotional resonance. Her work demonstrates self-awareness and profound critique while exploring feelings and identity concerns. Atwood's (1995) collection *Morning in the Burned House* is a powerfully personal work that discusses existential themes of loss, grief, memory, and individuality. Atwood's use of symbols in this collection resonates with Iraqi writer Al-Malaika's (1949) collection *Shathaya wa Ramad* (Shards and Ashes), which relies on symbolism to discuss traumatic memory and emotional and existential anxieties. Al-Malaika is known for her contributions to modern Arabic poetry, with a style marked by innovation, social critique, and an authentic personal voice that breaks with gender expectations.

Through a comparative reading of their poetry, this paper argues that symbolism functions as an instrument for expressing post-traumatic memory and reshaping the self. The collections reveal struggles with distortion and fragmentation while attempting to escape psychological entrapment. Therefore, the psychological approach fits for

investigating the crucial impact of traumatic memory on the poetic representations of Atwood and Al-Malaika. Selected poems are analyzed from a comparative psychological perspective to answer the following questions. How does symbolism function as a medium for articulating post-traumatic memory in the selected works? How does poetic language enable the reconstruction of the self after trauma? What similarities and divergences appear in Atwood's and Al-Malaika's use of symbols in representing post-traumatic memory and identity?

II. LITERATURE REVIEW

A. *Traumatic Studies and Memory*

Post-traumatic memory has received considerable attention from literary communities, which have explored the connections between trauma, memory, and post-experience representations through philosophical and psychological perspectives. From a psycho-scientific perspective, Clewett and Murty (2019) explained the profound influence of negative emotional experiences on memory and the significance of memory retrieval, recollection, and encoding in the process of reliving trauma. Based on experiments with trauma victims and random participants, they argued that emotional memories were more vivid when they are aroused. This process contributes to the recurring sensory pain associated with post-traumatic experience. Through the lens of Lacanian theory, LaCapra (2016) explained the role of traumatic memory in the construction of self-image and collective identity. Using examples of major historical traumas, such as the Holocaust, the U.S. displacement of American Indians, and the 9/11 attacks, LaCapra explored the lasting effects of trauma on the language that constructs victims' collective memory. Through literary texts, he also demonstrated the function of archive and testimony in oral history and collective memory while passing on and preserving historical traumas.

Alhourani et al. (2025) asserted that trauma is a layered experience of interrelated forces as gender, race, poverty, and social exclusion. It was suggested that trauma is not only confined to personal memory but integrated within structures of power and social meanings. The fragmentation of memory and the difficulty of articulating painful experiences demonstrate the psychological impact of trauma.

Van der Kolk (2002) investigated the common assumptions concerning memory and post-traumatic stress disorder, including formation, recollection, and the fragmentation and unreliability of traumatic memories in the process of deformed retrieval as a self-defense mechanism. Misperceiving these traumas has a strong effect on whether they are translated accurately into a narrative. People respond to trauma in diverse ways, connected to different biological factors and nervous readiness that sometimes leads to a misrepresentation of the self and others (van der Kolk, 1994).

Caruth (1996) argued that the primary source of pain in trauma emerges from the confusion that often coincides with traumatic memories. Relying on the theories of Freud, Lacan, Kant, and others, she explained the connection between public traumas and the formation of history. Following Freud's approach in defining trauma, Caruth used literary texts to explain the agony of trauma retrieval through memories. She emphasized the heavy influence of these memories during the healing process on the self, body, and time and explored the function of language in fathoming a trauma while representing it.

While these psychological and theoretical studies primarily addressed historical and cultural dimensions of trauma, they further provided a framework for understanding the difficulties in representing and articulating traumatic memories. This concern is pivotal to the present analysis of the selected poetic works of Atwood and Al-Malaika.

B. *Literary Representations and Trauma*

The variety of traumatic experiences, whether personal or public, leads to a wide range of representations of them. Whitehead (2004) examined themes and styles of distinguished modern authors, such as Toni Morrison, Pat Barker, and Anne Michaels, where a painful memory can haunt characters in different ways. Intertextuality and repetition in literary texts discussing traumatic memories can create a stark space and temporal consciousness within transgenerational memory. Probing into two novels of Kazuo Ishiguro, Trimarco (2023) argued that public trauma is ethical and collective, while its traumatic recollections are disruptive and difficult to fully recover or explain.

In *Trauma in Contemporary Literature*, post-traumatic representations in major literary works are introduced and analyzed. Marita Nadal and Monica Calvo argue in this book that trauma writing has layers of functions: a genuine process of healing and recovery, a historical documentation of private and public silenced experiences, and an ethical responsibility that encompasses the victim and the witnesses. While verbalizing trauma appears to be an effective means of resistance and healing, Kostova and Martinez-Alfaro admit the difficulties in rendering these experiences linguistically without resisting the tendency towards fragmentary language and repetitive narration.

Articulating certain experiences through figurative language may intensify meaning and add other dimensions to these experiences. The symbols of fire and burning have long been rooted in human experience, deeply connected to our imagination, memory, and personal reflection of the past (Bachelard, 1964). Fire can be seen as a poetic force and intimate symbol, because it fuels creativity and mirrors human desires and fears. Between immortality, purification, and destruction, fire was given a divine presence and power in Ancient Greek songs (Stuligrosz, 2020). Fire is a conceptual image connected to readers' emotions and passion (Karlson-Weimann, 2014). It is always contemporary since its

subjective connotations are related to history and culture. Therefore, fire is metaphorical before it becomes physical. It achieves its importance through the sense of estrangement that readers encounter in different texts.

Like fire, ashes have come to carry various meanings. Marder (2012) conceptually categorized fire into the sun of reality and the fire of the cave. These are fires of spirit, revolution, and ideology. People in recent times are dealing with the ashes (remains) of these two extinguished fires, symbolizing fragmentation, victimhood, unclassifiable society, and resistance. On another level, ashes have been discussed as a symbol of protection, ritual closure, transformation, and renewal (Adams & Roth, 2019). In this sense, ashes have a transformative meaning as a transformation of fire. They symbolically create social memory and identity over time and mark transitions like life and death.

Previous studies focused on the literary representations of trauma in prose to discuss linguistic resisting strategies against post-traumatic memory, which left an underexplored area in poetic narration. This study re-examined the significant symbols and metaphysical concepts in the poetry of Atwood and Al-Malaika as manifestations of post-traumatic memory and psychological anguishes. It sought to explain how individuals with post-traumatic memory experience the present through the lens of distorted memories.

C. Studies on Atwood and Al-Malaika

Atwood's (1995) *Morning in the Burned House* has received considerable critical attention. Pylvainen (1996) characterized the development of Atwood's awareness and perception throughout its five chapters as a phoenix rising from the ashes. What starts as a physical relationship with the self and the other leads to a conscious spirituality of selfhood. Conversely, Jamieson (1993) presented another reading, arguing that in this collection, Atwood resides in a liminal space between life and death due to grief over her dead father. Jamieson suggested that Atwood's struggle with pain and tension is a result of her being caught between a need for fulfilling societal expectations and a deep urge for fathoming her angst.

Lucas (2006) agreed that Atwood is entrapped within a temporal and metaphysical liminality. According to her, the traumatic experience of loss and absence in *Morning in the Burned House* produces a fragmentary language and a partial memory, leading to an opened closure that resists final recovery. In that work, the symbolism of fire, with its transformative meaning, lies in defying the decay of time and enduring powerful emotions and memories.

Naadiya Yaqoob Mir (2023) argued that Atwood's fiction presents traumatic women as victims of patriarchal violence, social constraint, and personal rupture. She further emphasized the function of memory, embodiment, and fragmented structure in the discourse of Atwood's female protagonists. Through the lens of trauma studies, Nazki (2024) asserted that Atwood's early novel *The Edible Woman* is considered her foundational literary work for exploring women's trauma and identity formation. It reveals her interest in the fractured female ego under patriarchal pressure.

In contrast to *Morning in the Burned House*, Al-Malaika's (1949) *Shathaya wa Ramad* has received far less attention beyond the scope of its contribution to modern Arabic poetry. Al-Ta'i (2016) explored how titles structurally and symbolically function in this work, connecting fragmentation, displacement, contrasting, and intertextuality to the traumatic emotional experience of the poet. He argued that the poet's use of symbols in titles, such as ashes and fragments, creates a transparent ambiguity and a linguistic transformation in readers' minds. He asserted that this process can enrich meaning and encourage contemplation and engagement with emotional experiences.

According to Al-Tuhami and Amin (2022), symbols in Al-Malaika's collection manifest the dark side of her life: her breakdowns, disappointments, and traumas. Through the dualism of light and darkness, her poetry reveals a profound need and persistent search for recovery and peace. Likewise, Muarich (2025) connected the pessimistic tone, the sense of grief, and the apparent alienation of Al-Malaika to her sensitivity to suffering and psychological angst. These feelings of sorrow and despair, according to Muarich, contributed to the creativity of Al-Malaika's poetry.

Nazar (2021) explained the role of Al-Malaika's poetry in exposing the psychological and social struggles of Arab women. He argued that her poetry exposes a profound sense of inner struggle. This struggle, according to Nazar, was generated by a conflict between traditional expectations and Arab women's self-awareness. Moreover, Nahhay (2020) suggested that a combination of private and collective experiences contributed to the sorrowful tone and fragmentation in Al-Malaika's poetry. He examined the psychological and socio-historical forces that shaped her poetic identity and influenced her engagement with pain.

Despite the informative insights of the existing studies on the works of Atwood and Al-Malaika, the role of post-traumatic memory as a psychological force in shaping their poetic language remained unexamined. The current study offered a comparative analysis of their foundational poetic works to explain how traumatic memory is articulated and represented in different cultural contexts.

III. METHODOLOGY

This study adopted a qualitative, text-based psychoanalytical approach to examine the representation of traumatic memory and its effects on poetry. The analysis focused on selected poems from Atwood's (1995) *Morning in the Burned House* and Al-Malaika's (1949) *Shathaya wa Ramad*. The procedures involved a close reading of each poet's texts to investigate and elucidate the poetic language. Guided by trauma studies, careful attention was paid to poetic representations, symbolic language, imagery, fragmentation, and emotional articulation as expressive mechanisms of traumatic memory. The analysis traced the psychological changes and personal development across the selected texts to demonstrate how individuation is expressed symbolically and poetically. The thematic focus of this study centered on the

poets' treatments of perceptual emotions within the metaphysical entrapment of post-traumatic experiences. After individual analysis, the works were compared to identify convergences and divergences.

This study was limited to the poets' experiences with traumatic memory, related concepts, and figurative language. Due to the sensitive nature of the poets' experiences of trauma, the methodological mood of this study was ethically attentive and reflective, remaining mindful of emotional complexity and nuanced psychological readings of the poetic expressions. Moreover, this study was limited by the lack of a published English translation of Al-Malaika's (1949) work, which was originally written in Arabic. Hence, all translations of her poems are my own, which may influence the interpretation of her work.

IV. DISCUSSION

A. *Morning in the Burned House*

In *Morning in the Burned House*, Atwood (1995) explores post-traumatic memories in a poetic narration. In the poems, her childhood memories are distorted and surreal. There is no indication of the nature or date of her trauma, but it still influences her perception of her memories. For her, the traumatic memory has a color and a presence, as can be seen in the poem "Waiting." Linked to her childhood, the memory is aged, nostalgic, and poignantly fragile like yellowed paper. She describes expecting it to be "a dark thing" that "would hide / in your closet" waiting to strike (p. 8) but instead finds it to be "only a memory":

And now it is now
and the dark thing is here,
and after all it is nothing new;
it is only a memory, after all:
a memory of fear,
a yellowing paper child's fear. (p. 10)

The first poem in the collection, "You Come Back," is a preface to intrusive childhood memories that take place throughout the collection between the family house and the woods. In this poem, Atwood (1995) shows a scene of perplexed thoughts in the old house, where questions like "What's been going on / while I was away? Who / got these sheets dirty, and why / are there no more grapefruit?" (p. 3) imply the elapse of time. However, she later explains the situation: "You know it was you / who slept, who ate here, though you don't / believe it" (p. 3), emphasizing a painful awareness of an entrapment. Memories, consequently, transcend temporal boundaries and persist into the present. Freud (1961) named the process of reliving the past the repetition compulsion. Instead of avoiding a traumatic memory, the sufferer unintentionally repeats certain behavior or re-experiences the repressed painful memories in the unconscious since they are not fully processed or fathomed (p. 14).

Spatial and temporal entrapments are symptoms of this traumatic memory in Atwood's (1995) work, simultaneously presented in symbols of burning and ashes. These symbols are linked to both the experience of transformation and her parents. Across the collection, the mention of burning coincides with change and transformation. Even though it is not an actual physical burning, it has clearly left a profound influence on her psyche and perception of her surroundings. Burning is a painful transformative process. Sometimes, it is meant to create new life from the ashes, as shown in "A Fire Place," which describes how "Earth does such things / to itself: furrowing, cracking apart, bursting / into flame" and replacing something with "scrubland, a light-green / sticky new forest" (p. 116). Such destruction can thus be seen in a positive light. This perspective contrasts with how people often "regret / the perishing of the burned place. / Only we could call it a wound" (p. 117). They struggle with painful memories that expand in time and kindle inner flames. The fire within influences the psyche and transforms the identity as a survival mechanism to cope with difficult situations.

Throughout her memories in nature, her father's presence is central to the process of burning, due to his responsibility for kindling the fire. The "father chops with his axe" while "His gun leans behind the door" and "Smoke comes out of the metal chimney" (Atwood, 1995, p. 76). The main representative objects of her father inside the house are the fire, his gun, and his boots. In the woods, she remembers herself "Holding the log / while he sawed it" (p. 91). His emotional absence in her childhood memories is represented by him being busy in the woods. More recently, she has an auditory memory: "It's nineteen ninety-four, / I can hear the sound of the chopping" (p. 77). In her father's final illness, he remembers in fragments and likes to "talk about axes" and "the many names of wood" (p. 89). He says that "We need more wood" because "winter's on its way" (p. 84), as a necessity and duty that must be fulfilled. However, descriptions of his role as a father figure lack an emotional presence, signified, for example, by her description of "looking hard and up close at the small / details" such as "the blackish and then the greying / bristles on the back of his neck" (p. 91) and "his square finger, earth under / the nail" (p. 92), investigating his materiality. Such details suggest he did not render warmth and affection as a provider of fire but rather heat, distance, and burning.

Morning memories in Atwood's (1995) burned childhood house are vividly persistent, almost surreal in their imagery. In the poem "Up," she wakes up paralyzed in bed and again refers to burning to encourage herself to get up: "Pretend the house is on fire / and you must run or burn" (p. 111), but she is held down by memories:

No. Nothing so simple. The past, its density
and drowned events pressing you down,

like sea water, like gelatin
filling your lungs instead of air. (p. 110)

In this sensory-heavy image, Atwood (1995) is suffocated by the sluggish movement of submerged events that penetrate into her psyche and create an emotional burden. These unresolved memories hold her back and prevent her from moving forward into the future.

In the final poem, which shares the name of the collection ("Morning in the Burned House"), Atwood's (1995) childhood house is burned down, but she still sees herself alone inside, with a lucid depiction of the objects around her. The image of a "spoon which was melted scrapes against / the bowl which was melted also" (p. 126) reminds of Salvador Dalí's famous painting *The Persistence of Memory*. Similar to the melting clocks in that work, objects in Atwood's burned house lose their fixed shapes, implying her subjective experience of time due to a traumatic memory. Scattered around absurdly, the "clothes are still on the hangers," and there is a "tin cup and rippled mirror" (p. 126), representing distorted memories that are pointless and difficult to perceive and explain. This follows from the tendency where "memories of trauma may have been organized on an implicit or perceptual level, without an accompanying narrative about what happened" (van der Kolk, 2002, p. 12).

While symbols of fire and burning are linked to Atwood's (1995) father, ashes are associated with her mother. As a symbol, ashes represent what remains: loss, memories, and ruin. They can also be regarded as signifying resilience and hope for a new beginning. Her mother in the old house patiently "rakes the ashes" (p. 76) after dead fires to collect the remains and hope for a resurrection. This action is her persistent female resilience in a materialistic masculine environment. The smell of ashes in the house is pervasive, as "The carpet smells of ashes" (p. 77). The olfactory information connected to her mother contrasts with the previous auditory description of her father chopping wood. Both are profoundly sensory in childhood memories, but the olfactory content is more emotional, vivid, and evokes nostalgia. This fragmentary expression of memories follows from a description of traumatic memories by van der Kolk (2002):

Memories of the trauma tend to be experienced as fragments of the sensory components of the event, as visual images, olfactory, auditory, or kinesthetic sensations, or intense waves of feelings that patients usually claim to be representations of elements of the original traumatic event. (p. 14)

The old house is connected to Atwood's (1995) mother and certain familial activities. A paradoxical memory when "listening to the radio, news of disasters / that made you feel safe, / like the voice of your mother" (p. 9) carries a layered meaning of her psychological state: feeling comfort while hearing news of pain and fear. This detail implies that she was raised in chaotic times, when what should be anxiety-inducing news could be soothing in the family house. In another memory of the house, an atmosphere of pervasive peace and comfort takes over in 1943, when "children dance around" a bonfire, "singing about the" Second World War (p. 76). In this memory, this war "is happening elsewhere," yet she later retracts this notion, emphasizing that everything in her memory "recurs / and nothing is elsewhere" (p. 77). According to the conscious version of her mature self, what she once perceived as remote anguish actually exists within her.

Across the collection, dwelling in the past intertwines with a new distorted perception of the self and the body. In many of her memory poems, Atwood (1995) regularly uses the pronoun "you" as if conversing with the reader for the sake of engagement. In fact, by insisting on using this pronoun, she merely reveals the presence of another version of herself, which continuously relives the past in the old house: "someone else / has been here wearing / your clothes and saying / words for you" (p. 3). This is a motif she uses in reflective poems to emphasize a fragmented self that failed to escape a traumatic memory. According to van der Kolk (2002), "certain happenings ... leave indelible and distressing memories--to which the sufferer continually returns, and by which he is tormented by day and by night" (p. 6).

In these poems, self-transformation into different forms occurs in phases and is associated with a developing perception. Atwood (1995) justifies and expounds on the first transformation through the Greek myth of Daphne. Like Daphne, she escapes into another form, hiding from the anonymous man, who is the only one who knows what happened and witnessed her transformation: "before bark/fur/snow closed over / my mouth, before my eyes grew eyes" (p. 26). Her new form is silent, yet perceives more. While Daphne transformed into a harmless laurel tree, she transformed into a being with "eight fingers / and a shell, and live[s] in corners" (p. 27), working on her ideas: "venom, a web, a hat" (p. 27). She lost her sense of the wholeness with time and replaced it with a fragmented perception of her body: "When I was all body I was lazy. [...] Now there are more of me" (p. 23). During her entrapment, she looks around in the burned house and observes missing parts of her body: "I can't see my own arms and legs" (p. 127). Her physical form is no longer whole or visible, as if she transcends through her memory into another ethereal form inside her "burning clothes" that are "holding my cindery, non-existent, / radiant flesh. Incandescent" (p. 127).

Her awareness of herself as layered grows deeper, as shown in "The Signer": "in an area of darkness behind my head / stands a woman dressed in black, / even the stockings: my unknown twin" (Atwood, 1995, p. 114). This woman who is mourning and trapped in her traumatic memory could represent her Shadow self. According to Jung (1969), "Everyone carries a Shadow, and the less it is embodied in the individual's conscious life, the blacker and denser it is [...] if it is repressed and isolated from consciousness, it never gets corrected, and is liable to burst forth suddenly in a moment of unawareness" (p. 77). Since the poet "speak[s] / in a kind of blindness" (Atwood, 1995, p. 114), the woman in black takes the stumbled words and turns them into something "solid, / [that] become a gesture, a skein, / a semaphore of the body / for those who listen with their eyes" (p. 114). Bringing the Shadow into consciousness creates a sense of profound introspection and a perfect duality: brightness and darkness, fragility and strength, and blindness and discernment. Even

though her Shadow is taking control of her words in this phase, realizing its presence and accepting its influence on her life suggests that she is heading toward individuation. Jung considered this process of discovering and accepting the different sides of oneself as a means of psychological healing and finding one's true self (as cited in Dobie, 2012, p. 63).

B. *Shathaya wa Ramad*

Post-traumatic memory in Al-Malaika's (1949) *Shathaya wa Ramad* is articulated in different ways, gradually from a melancholic and desperate tone into a sublime voice that transcends temporal liminality. Through her poems, Al-Malaika expresses her deep sorrow and resentment toward her emotional trauma, which was probably recent based on the intense feelings expressed. She treats concepts in her poems as beings and gives them human traits for them to be part of her emotional experience. Figurative language symbolizes her psychological state in a poetic reflection of an unacceptable situation. The title of the collection *Shathaya wa Ramad* (Shards and Ashes) symbolizes her experience as a trauma survivor and its ramifications. The word *shathaya* (shards) suggests the poet's inner anguish and emotional wounds. The haunting of painful memories turns her life into fragments of struggle that are reflected in her psyche. The word "shards" resonates with the cultural memory of post-war consciousness. Although the term "ashes" evokes feelings of loss and regret, it also stands for renewal, resurrection, and hope. It is a sign of the poet's resistance and the gradual healing in the aftermath of trauma.

As a modern poet, Al-Malaika heavily relies on metaphors and symbols to intensify meaning and reflect the layers of psychological experience. In "Tawārīkh Qadīmah wa Jadīdah" (Old and New Histories), she starts a recurrent metaphor; the past is portrayed as an entity that lives, dies, is buried in darkness, and lives again. In this poem, she initiates a metaphysical journey that extends throughout the collection to find and revive the past (Al-Malaika, 1949, p. 47). Since trauma does not make sense to the survivor, the external world fails to provide a conceivable reality, and a dissociation occurs. Therefore, she looks inside to construct meaning out of painful memories. As van der Kolk (2002) noted, "It is precisely because there is no immediate accommodation that there is complete dissociation of the inner activity from the external world. As the external world is solely represented by images, it is assimilated without resistance to the unconscious ego" (p. 6). Al-Malaika (1949) relies on metaphors and surreal imagery to draw an alternative inner landscape. In this deeply internal quest into the unconscious, she seeks to comprehend her anguish, fears, and desires. Through her memories, she sees fragments of dysfunctional body parts that have certain important connotations in Arab culture. The forehead, which normally signifies dignity and honor, is blind. The eyes, which indicate knowledge and insightfulness, are silent. The heart, which provides affection and spirituality, is barely a remnant. All these fragments imply loneliness, injustice, and lack of support in the past. While the dead past is buried in a forsaken coffin, the future drags its half-paralyzed body, indicating extended suffering, "and tomorrow life will grow / upon the painful wound" (p. 49). It is a cumulative experience of painful memories that will not allow the future to thrive and prosper.

Al-Malaika's (1949) poetry does not reflect the external world alone; it exposes the fragmentary internal self as well. In "'Indamā Unbi'itha al-Māḍī" (When the Past Is Resurrected), she hears voices in the dark coming from the past. Facing painful memories, she "stumbled upon the remains of [her] youth" (p. 57). The pain is so acute that her youth is scattered by the sharpness of her memories. Every time the memory reaches her with its hands, the poet's soul and heart are ripped off by her nails (p. 58), an attempt to escape and overcome the pain. In "'Urūq khāmidah" (Dormant Veins), she perceives herself as a distorted illusion: "Our eyes with no color / to reflect things," "Our lips are mere meat [and] lack of tones," and "our fingers are dead / with no depth" (p. 67). Because of such a traumatic memory, she is stuck with dysfunctional organs and a vaguely defined identity. Her senses cannot provide the certainty needed to eliminate doubt.

Throughout this collection, traumatic memory is figuratively illustrated with psychological implications. In "Al-Af'uwān" (The Serpent), her memory is portrayed as a serpent that comes after her, crossing temporal and spatial boundaries: "behind the translucent fog / lurks a gigantic serpent / I cannot escape" (Al-Malaika, 1949, p. 79). The symbolism in the title suggests the influence of Mesopotamian mythology, in which the serpent represents the primal chaos, as embodied by the goddess Tiamat as described in the introduction to *Enuma Elish* (Dalley, 2000, pp. 228-230). In the *Epic of Gilgamesh*, the serpent reflects the reality of human limitations, a creature that must be defeated or it will ruin people's lives (Dalley, 2000, pp. 151-153). Al-Malaika's foe comes from within, from her unconscious. It chases her through darkness and mocks her screams and silence. Therefore, her attempts to escape are futile while she is burdened with dark thoughts and doubt:

The farther I flee
the higher my steps transcend the summits
yet what daylight's labor had undone
returned with chains of memory
I shall not desire release
for what release is there
while my terrifying foe
his eyes spit autumn
over a soul that yearns for spring (p. 78)

Recognizing the reality of her "foe" and accepting its existence contribute to her rising self-awareness. Instead of being pulled down by the heaviness of sad memories, she climbs above her fears and anger. The contrasting image of autumn and spring indicates the liminality that she intends to escape from and her desire to move forward.

The symbols of fire and burning in this collection represent her resilience in the face of a traumatic experience. In “Al-Jurh Al-Ghāḍib” (The Furious Wound), her repressed anger from unjust experiences agitates an internal fire: “My silence turns into a fire screaming across the horizon” (Al-Malaika, 1949, p. 70). Her world crumbles in this poem, the previously paralyzed future is now a fragmented corpse, and the dreams fall on earth as fallen stars (p. 71). The internal fire unleashes a hidden part of her that refuses to be repressed. This could be understood from Jung’s (1969) assertion:

We carry our past with us [...] and it is only with an enormous effort that we can detach ourselves from this burden. If it comes to a neurosis, we invariably have to deal with a considerably intensified shadow. And if such a person wants to be cured it is necessary to find a way in which his conscious personality and his shadow can live together. (p. 77)

Since Al-Malaika’s (1949) ego suffers in silence, her shadow rebels against the stereotype of the submissive female. Her inner world integrates with the outside natural world in response to her frustrated self. Hence, the night, the moon, stars, waves, and darkened clouds become angry with her (Al-Malaika, 1949, p. 70). This cosmic sympathy indicates a transcendent self, where the intense feelings resonate with the natural world. The state of introspection associated with her internal burning participates in her journey of self-realization. While remembering traumatic memories of her life, she tosses out reflective questions, such as the following: “how to endure sadness while burning with thirsty tears?”, “how does the bitterness of despair turn into a burning fire inside?”, and “is there any past dream / to survive burning?” Any “unburned hopes?” (p. 134). Through these questions, she inspects her boundaries, endurance, and strength.

Al-Malaika’s (1949) attitude toward the ashes in this collection is contradictory, torn between hope and sorrow, and they represent an infinite resistance, as in “beware of this ash / within its depths living embers” and “you who thought fire is a dormant clay / and forgot the tempest of youth and its immortality” (p. 169). A cycle of burning and ashes continues till her soul is purified and elevated. She “screamed at the material earth / to release the exhausted youth” since it is burned with flames (p. 170). She intends to “break the chains with her hands” and “fly high from the past to the future” (p. 171), creating a new life from her painful memories, fears, and sighs. She realizes that internal fire brings life, and the burning lifts her above the material world. Nevertheless, Al-Malaika reveals in “Ramad” (Ashes) her sense of sorrow regarding life’s cruelty (p. 182). Ashes in this poem represent the pain of loss. She investigates the remains (ashes, a faded planet, ancient stories, silence, empty glasses, and painful memories) and comes to the conclusion that “no image pulses with life / nothing but ashes” (p. 184). Even with this melancholic tone, she unconsciously presents ashes as a transformative phase: “we are still shoveling the ashes / to feed the stove / and the hands of hope vainly / pray for tomorrow” (p. 183). Ashes in this work are a multilayered symbol that conveys contrasting meanings between entrapment and emancipation, acceptance and resistance.

C. Atwood and Al-Malaika

Examining the representation of traumatic memory in Atwood’s (1995) *Morning in the Burned House* and Al-Malaika’s (1949) *Shathaya wa Ramad* revealed key similarities and differences between the two poets. Both demonstrate a painful psychological entrapment in the past that leaves them with a fragmented self and prevents them from moving on. Atwood (1995) explains this as follows: “What prevents you? The future. The future tense, / immense as outer space. / You could get lost there” (p. 110). Similarly, Al-Malaika (1949) describes her situation by asking, “where to escape from my enemy’s spiteful lashes / In my path they spit unbearably dead tomorrow?” (p. 79). Memories of the past destroy any hope of a promising future. Thus, both poets portray traumatic memories as horrible entities that haunt them for a long time. Although they narrate the ramifications of traumatic memories, they both assert the deficiency of their bodily senses regarding intense emotions, as mouths cannot speak, eyes cannot see, and hands cannot reach. According to Al-Malaika, “we are here illusions, no color / no voice, no shape. Mirage of nothingness, no meaning / no utterance no shade” (p. 66). The self-image is destroyed when the past is deformed due to painful memories and doubt.

Atwood (1995) relies on the material world to represent her memories while she pensively expresses her struggle with post-traumatic memories. Al-Malaika (1949), on the other hand, presents her experience with a tone of protest and questioning, which contributes to the chaotic atmosphere of her metaphysical world. This world that she creates for her quest mirrors her rejection of her cultural boundaries. In this world, abstract concepts are personified to express her true feelings and thoughts: “I rage for the trembling and complaining wound / With me the slaughtered quivering patience will go mad” (p. 72).

Both poets use burning and ashes as central images of post-traumatic transformation, but their perspectives vary according to their different focuses. Atwood’s (1995) recurrent use of burning reflects her desire to comprehend the origins of her trauma: “Who is it, exactly, you have needed / all these years to forgive?” (p. 111). A question needs to be answered and healed before moving on. On the other hand, Al-Malaika (1949) asks questions to fathom the consequences of her trauma: “this is me no doubt / why can I not touch it with my hands? / Why can I not touch myself? / And end my eternal burning?” (p. 165). At the same time, for both poets, burning symbolizes a ritual of transformation and transcendence. The end of each collection emphasizes an internal growth, an acceptance of their inner realities, and symptoms of psychological individuation. Atwood (1995) asserts the importance of accepting oneself: “together we are practicing / for the place where all the languages / will be finalized and / one; and the hands also” (p. 115). With her shadow, they form one true self that is capable of surviving and resisting.

Although ashes can be regarded as a symbol of death and loss, the psychological development the poets reveal by the end of each collection suggests that ashes stand for resistance in *Morning in the Burned House* and for the liminal space

before resurrection in *Shathaya wa Ramad*. Al-Malaika (1949) pessimistically states that “our ghosts carve their way / through the trembling graveyard of memories / no image pulses with life / nothing remains but ashes” (p. 184). The confident tone at the end of the collection never denies her awareness of the damage within. Therefore, she relies on ashes to symbolize the irreversible outcomes of her emotional experience.

V. CONCLUSION

This paper offers a comparative psychological analysis of two poetry collections: *Morning in the Burned House* by Atwood (1995) and *Shathaya wa Ramad* by Al-Malaika (1949). It brings forth the role of symbolism in translating painful experiences that resist direct narration. The poetic imagery embodied in the symbols of fire and ashes facilitates the emotional demonstration while preserving its intensity within certain cultural contexts. Drawing on the transformative symbolism of fire and ashes, both poets depict the post-traumatic memories while simultaneously illuminating the psychological self-transformation. By offering a resilient, symbolic mode of expression, poetic language defies traumatic disruption and allows self-reconstruction after trauma. This procedure endowed their poetic voices with scope and identity within the broader context of women’s poetry. Although the two poets mainly rely on fire and ashes to represent post-traumatic memory, the comparison reveals a notable divergence in their poetic lenses. Since Atwood’s traumatic memory is connected to her parents and childhood, her focus is pensively psychological. This focus allows her to reach individuation early in the process of self-reconstruction. Al-Malaika’s symbolism is closely tied to a recent, socio-cultural experience of her life, which accounts for her vague and stark emotional imagery. This comparative analysis deepens the understanding of traumatic memory in women’s poetry by highlighting the works of two modern female poets with different cultural backgrounds. The study emphasizes the role of poetry as a means of survival and self-reconstruction. It paves the way for further interdisciplinary studies in literary representations of trauma, and calls for further research in the maturation of women’s poetry and its pivotal role in shaping both identity and poetic voice.

ACKNOWLEDGEMENTS

This work was funded and supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [KFU260675].

REFERENCES

- [1] Adams, E. C., & Roth, B. J. (2019). Introduction. In E. C. Adams & B. J. Roth (Eds.), *Agent of change: The deposition and manipulation of ash in the past* (pp. 1–12). Berghahn Books.
- [2] Armstrong, C. I. (2020). Trauma and poetry. In C. Davis & H. Meretoja (Eds.), *The Routledge companion to literature and trauma* (pp. 540–555). Routledge.
- [3] Alhourani, M. I., Elhalafawy, A. I., Zoair, Z. M., Khalifa, S. H., Abou Adel, M. A., & El-Naggar, I. A. (2025). Beyond the mother knot: Trauma and intersectionality in Suzan-Lori Parks’ *In the Blood*. *Theory and Practice in Language Studies*, 15(8), 2486–2495.
- [4] Al-Malaika, N. (1949). *Shathaya wa ramad* [Fragments and ashes]. Dār al-‘Awda.
- [5] Al-Ta’i, L. Y. H. (2016). The semiotics of the title in *Shadhaya wa Ramad* by the poet Nazik Al-Malaika. *Midad Al-Adab*, 12(1), 144–156. Retrieved December 22, 2025, from <https://digitalcommons.aaru.edu.jo/midad/vol12/iss1/6>
- [6] Al-Tuhami, N. H. M., & Amin, H. A. K. M. (2022). Symbolism in the poetry of Nazik Al-Malaika. *Journal of Linguistic and Literary Studies*, 13(2), 170–186. Retrieved December 22, 2025, from <https://journals.iium.edu.my/arabiclang/index.php/jlls/article/view/999>
- [7] Atwood, M. (1995). *Morning in the burned house*. HarperCollins.
- [8] Bachelard, G. (1964). *The psychoanalysis of fire* (A. C. M. Ross, Trans.). Routledge & Kegan Paul. (Original work published 1938)
- [9] Caruth, C. (1996). *Unclaimed experience: Trauma, narrative, and history*. Johns Hopkins University Press.
- [10] Clewett, D., & Murty, V. (2019). Echoes of emotions past: How neuromodulators determine what we recollect. *eNeuro*, 6(2). <https://doi.org/10.1523/ENEURO.0108-18.2019>
- [11] Dalley, S. (2000). *Myths from Mesopotamia: Creation, the flood, Gilgamesh, and others*. Oxford University Press.
- [12] Dobie, A. B. (2012). *Theory into practice: An introduction to literary criticism* (3rd ed.). Wadsworth Cengage Learning.
- [13] Freud, S. (1961). *Beyond the pleasure principle* (J. Strachey, Trans.). The Hogarth Press. (Original work published 1920)
- [14] Jamieson, S. (1993). Mourning in the burned house: Margaret Atwood and the modern elegy. *Canadian Poetry*, 38, 38–57.
- [15] Jung, C. G. (1969). *The collected works of C. G. Jung* (2nd ed., Vol. 11, R. F. C. Hull, Trans.). Princeton University Press.
- [16] Karlson-Weimann, D. (2014). *Burning images: The metaphor of fire in literature* [Bachelor’s thesis, University of Gothenburg]. Retrieved December 22, 2025, from <http://hdl.handle.net/2077/38182>
- [17] LaCapra, D. (2016). Trauma, history, memory, identity: What remains? *History and Theory*, 55(3), 375–400. <https://doi.org/10.1111/hith.10817>
- [18] Lucas, R. (2006). Incandescence: “The power of what is not there” in Margaret Atwood’s *Morning in the Burned House*. In J. Moss & T. Kozakewich (Eds.), *Margaret Atwood: The open eye* (pp. 319–330). University of Ottawa Press.
- [19] Marder, M. (2012). After the fire: The politics of ashes. *Telos*, 161, 163–180. <https://doi.org/10.3817/1212161163>
- [20] Mir, N. Y. (2023). *Women and trauma in the works of Margaret Atwood and Anita Desai*. Cambridge Scholars Publishing.
- [21] Muarich, M. Q. (2025). The pain of separation for the two poets (Nazik Al-Malaika, Stevie Smith): A comparative study. *Lark Journal*, 17(1), 995–1014, <https://doi.org/10.31185/lark.4039>

- [22] Nadal, M., & Calvo, M. (Eds.). (2014). *Trauma in contemporary literature: Narrative and representation*. Routledge.
- [23] Nahhay, M. H. (2023). Motives of Pain in Nazik Al-Mala'ika's Poetry. *Journal of Language Studies*, 3(2), 289–297. <https://doi.org/10.25130/jls.3.2.19>
- [24] Nazar, S. (2021). Nazik Al-Malaika: Her poetic themes and contribution towards the identity crisis of Arab woman. *Tahdhīb al-Afkār*, 8(2), 7–18.
- [25] Nazki, S. H. (2024). Examining the female ego and trauma in Margaret Atwood's *The Edible Woman*. *Journal of Women Empowerment and Studies*, 4(4), 25–35.
- [26] Pylvainen, T. (1996). *Dawn of discovery: Margaret Atwood's Morning in the Burned House* [Master's thesis, Lakehead University]. Retrieved December 22, 2025, from <http://knowledgecommons.lakeheadu.ca/handle/2453/2478>
- [27] Stuligrosz, M. (2020). The imagery and symbolism of fire in the Homeric hymns and in Greek choral lyric. *Eos*, 107, 25–43. <https://doi.org/10.34616/e.2020.25.43>
- [28] Trimarco, P. (2023). The extremities of literature: Traumatic memory in two novels by Kazuo Ishiguro. *Text Matters: A Journal of Literature, Theory and Culture*, 13, 195–209.
- [29] Van der Kolk, B. A. (1994). The body keeps the score: Memory and the evolving psychobiology of post-traumatic stress. *Harvard Review of Psychiatry*, 1(5), 253–265.
- [30] Van der Kolk, B. A. (2002). Trauma and memory. *Psychiatry and Clinical Neurosciences*, 52(suppl. S1), 52–64. <https://doi.org/10.1046/j.1440-1819.1998.0520s5S97.x>
- [31] Whitehead, A. (2004). *Trauma fiction*. Edinburgh University Press.

Fatimah A. Alwadaani is currently an assistant professor of English Literature in the Department of English Language, King Faisal University, Al Ahsa, Saudi Arabia. She received her PhD in English Literature from Imam Mohammad ibn Saud Islamic University in June 2022. Her academic interest areas are poetry, cultural studies, literary theory, philosophy, and comparative literature. E-mail: falwadaani@kfu.edu.sa

Attitudinal Resources in the English Textbook *Life – Upper Intermediate*: A Systemic Functional and Intercultural Perspective

Ngoc T. B. Han

English Faculty, Thuongmai University, Hanoi, Vietnam

Abstract—This study investigates attitudinal resources in the international English as a Foreign Language (EFL) textbook *Life – Upper Intermediate* through the lens of the Appraisal framework (Martin & White, 2005). It focuses on how evaluative meanings are distributed and realized across different text types to shape stance and interpersonal positioning in pedagogical discourse. A qualitative discourse analysis was conducted on selected reading passages to identify instances of Affect, Judgement, and Appreciation, with particular attention to their polarity, distinguishing between positive and negative instances. The findings reveal that Appreciation predominates, highlighting the textbook's emphasis on valuing human achievements, social harmony, and cross-cultural understanding. Judgement and Affect appear more selectively, constructing moral perspectives and empathy in the narratives. These attitudinal patterns demonstrate how the textbook constructs an implicitly positive worldview aligned with global educational discourse. The study contributes to the understanding of evaluation in instructional materials and offers insights into the linguistic realization of stance in international EFL publishing.

Index Terms—attitudinal resources, stance, EFL textbook, intercultural communication, cultural values

I. INTRODUCTION

Evaluation is an essential dimension of human communication, shaping how people construct stance, negotiate meaning, and express shared values. Within the Appraisal framework (Martin & White, 2005), evaluation is theorized as a system of linguistic choices through which speakers and writers express emotions, judge behaviours, and appreciate phenomena. These attitudinal resources – categorized as Affect, Judgement, and Appreciation – constitute the foundation of interpersonal meaning in discourse, linking linguistic expression to cultural and ideological positioning. They enable texts not only to inform but also to persuade, align, and socialize readers into particular worldviews.

While evaluation has been widely examined in political, media, and literary discourse (e.g., Coffin, 2006; Thompson & Hunston, 2000; Hyland, 2005; Hood, 2010; Geng & Wharton, 2016), its realization in educational materials – particularly international English as a Foreign Language (EFL) textbooks – remains underexplored. Yet textbooks are among the most powerful forms of practical discourse: they shape not only linguistic knowledge but also the attitudes, ideologies, and cultural perspectives of learners. Through lexical and grammatical choices, textbooks construct implicit systems of value that guide how readers perceive the world and their role within it. As Gray (2010) and Tomlinson (2012) have shown, global ELT materials do not merely teach English; they embody cultural narratives and moral stances that frame what is desirable, admirable, or universal. However, relatively few studies have systematically examined *how* these evaluative stances are linguistically constructed or *what kinds of worldviews* they promote. This gap in Appraisal-based textbook research motivates the present study.

Addressing this gap is crucial for several reasons. Theoretically, analyzing evaluative meanings in textbooks extends Appraisal research beyond spontaneous discourse into a designed pedagogical genre – a domain where linguistic, educational, and ideological functions intersect. It allows investigation of how Attitude operates in texts that are crafted to motivate, instruct, and model cultural values simultaneously. Pedagogically, such inquiry exposes the hidden curriculum embedded in international teaching materials: the value systems, moral orientations, and affective positions that accompany language input. For teachers, awareness of these meanings enables more reflective material adaptation; for learners, it provides a path toward critical literacy and intercultural awareness – skills central to English for Academic Purposes (EAP) and global citizenship education.

In higher education, particularly in non-Anglophone contexts such as Vietnam, international coursebooks like *Life – Upper Intermediate* are widely used across university programs. They function as linguistic resources, but also as discursive models that socialize students into particular ways of seeing, feeling, and valuing. Despite this influence, little empirical work has examined how these materials linguistically encode stance and intercultural values. Accordingly, this study investigates attitudinal resources in twelve reading passages (approximately 6,800 words) from *Life – Upper Intermediate* (Dummett et al., 2021). Drawing on Appraisal-based discourse analysis, this study examines how Affect, Judgement, and Appreciation are realized in *Life – Upper Intermediate*, focusing on polarity patterns and their role in constructing stance and intercultural meanings.

II. LITERATURE REVIEW

A. Theoretical Background

Appraisal theory, developed within Systemic Functional Linguistics, offers a systematic model for analyzing evaluative meaning (Martin & White, 2005). It distinguishes three interrelated subsystems: **Engagement** (resources for dialogic positioning), **Attitude** (Affect, Judgement, Appreciation), and **Graduation** (scaling of meanings). Figure 1 presents an overview of the Appraisal Framework as a whole.

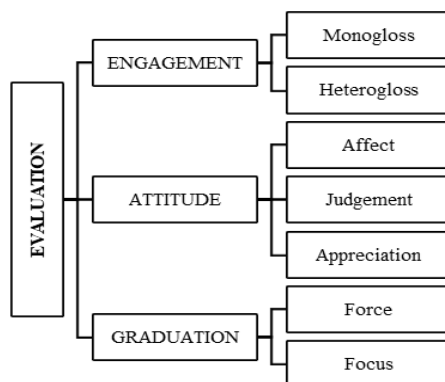


Figure 1. An Overview of the Appraisal Framework (Adapted From Martin & White, 2005)

In essence, Appraisal theory provides a comprehensive lens to explore how language encodes values and interpersonal meanings. Martin and White (2005) argue that the realization of Attitude tends to appear in all kinds of discourse, regardless of grammatical boundaries (p. 35). They therefore proposed a unified appraisal system comprising three key components: Engagement, Attitude, and Graduation. It explains how speakers and writers adopt stances, invite alignment, and scale their evaluations, thereby linking the micro-level of lexical choice with the macro-level of discourse and ideology. While the Appraisal framework comprises three subsystems, the present study focuses exclusively on Attitude, given its central role in encoding affective, ethical, and aesthetic values. Figure 2 illustrates the Attitude subsystem of the Appraisal framework, outlining its three domains – Affect, Judgement, and Appreciation – with their respective subcategories.

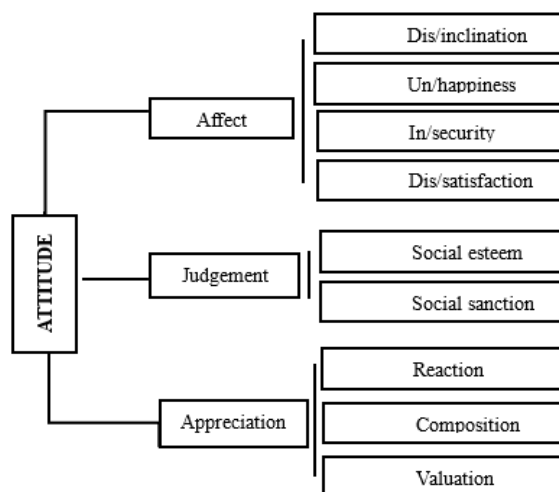


Figure 2. Attitude in Appraisal Framework (Adapted From Martin & White, 2005)

Affect is realized when emotional states are made explicit, as in “The captain felt happy (+ happiness) / sad (– happiness)”, which encodes feelings of joy or sorrow. Affect also covers desires and security, illustrated by “The student was eager to join the discussion (+ inclination) / reluctant to participate (– inclination)” and “She felt confident before the exam (+ security) / anxious about the outcome (– security)”. **Judgement** evaluates human behavior, capacity, or ethics. For instance, “He is honest (+ veracity) / dishonest (– veracity)” signals integrity, while “The leader was brave in facing challenges (+ tenacity) / cowardly in avoiding responsibility (– tenacity)” reflects perseverance. **Appreciation** refers to evaluations of objects, processes, and phenomena. This can be seen in “The book is remarkable (+ valuation) / dull (– reaction), The design was harmonious (+ composition) / unbalanced (– composition)”, or “The lecture was inspiring (+ reaction) / boring (– reaction)”.

These examples show how Attitude systematically encodes positive and negative values across emotions, behaviors,

and phenomena. They also highlight how evaluative meanings are not restricted to isolated words but operate through full sentences, shaping stance and alignment in discourse.

The theory builds on earlier and broader scholarship on evaluation. Thompson and Hunston (2000) and Channell (2000) both emphasized the interpersonal nature of evaluation, showing that it expresses attitudes and values to build solidarity between writer and reader and that corpus-based patterns of evaluative lexis shape positive or negative connotations in discourse. White (1998) had already extended Systemic Functional Linguistics to analyzing news discourse, revealing how journalistic language constructs ideologically informed judgements and laying the groundwork for later Appraisal studies. In his subsequent publications, White (2006, 2016) further demonstrated how journalistic and media texts use evaluative language to position readers. These studies – ranging from theoretical accounts (Thompson & Hunston, 2000), corpus-based insights (Channell, 2000), to discourse analyses of media texts (White, 1998, 2006) – demonstrate the diversity of approaches that underpin Appraisal theory. They all confirm that evaluation is not an optional stylistic feature but a fundamental dimension of communication, essential for constructing stance, aligning readers, and organizing discourse. Following these foundational studies, numerous subsequent works have also continued to apply Appraisal theory to a wide range of genres, further demonstrating its versatility in analyzing evaluative meaning. Studies have examined political speeches (Ross, 2016), advertising discourse (Sahragard & Meihami, 2016), and literary texts (Hashim, 2022). The extensive use of Appraisal theory across genres – from political and media commentary to academic writing and literary analysis – demonstrates its broad explanatory potential. The following section of the literature review narrows its focus to academic and educational discourse, where evaluation functions as a linguistic resource for constructing stance and intercultural values in EAP contexts.

B. Related Studies on Evaluation in Academic Contexts

Within EAP, evaluation has become a prominent concern, not only as a linguistic phenomenon but also as a pedagogical priority. Early studies established the foundation for understanding stance in academic writing, highlighting its interpersonal nature and revealing that non-native writers often underuse attitudinal resources, leading to more impersonal prose (Dontcheva-Navratilova, 2009). Subsequent large-scale and diachronic research, such as Hyland and Jiang (2016), further demonstrated how disciplinary and pedagogical contexts shape meta-discourse practices over time, underscoring the dynamic and context-sensitive nature of evaluative expression in EAP.

Recent research highlights both the continuity and innovation of Appraisal studies across academic, pedagogical, and technological domains. In higher education, evaluative language has been shown to support critical thinking and academic engagement (Chou, 2022), while also playing a key role in the construction of authorial voice through Graduation resources (Vo & Le, 2025). More broadly, stance is understood not as individual attitude expression but as a dialogic process through which writers negotiate knowledge, identity, and alignment with readers (Siromaa & Rauniomaa, 2021; Ebrahimi & Mohsenzadeh, 2022; Khamkhien, 2014; Kim, 2025). Across classroom and academic contexts, evaluative language consistently underpins disciplinary meaning-making – from legitimizing knowledge in science and history discourse (Martin & Veel, 1998; Coffin, 2006) to constructing stance, authority, and interaction in academic writing and feedback (Hyland, 2017; Hood, 2010; Geng & Wharton, 2016; Muti'ah & Azizah, 2024). Collectively, this body of research affirms evaluation as a central linguistic resource for constructing knowledge, argumentation, and engagement in EAP.

A promising yet relatively underexplored strand within Evaluation research in EAP concerns textbooks as carriers of evaluation, highlighting how instructional materials embed and transmit cultural, moral, and pedagogical values through language. Rather than neutral vehicles of information, textbooks construct meaning through evaluative framing that legitimizes particular values and perspectives. From Coffin's (2006) analysis of historical narratives to later studies on global and Asian EFL materials (Gray, 2010; Tomlinson, 2012; Myskow, 2017; Nguyen, 2023; Nguyen et al., 2025), evidence consistently shows that evaluative resources – especially Judgement and Appreciation – function to integrate language learning with ideological and moral education. These studies underscore that textbooks act as sites of socialization, embedding cultural identities and value systems within their linguistic fabric.

Despite growing attention to ideology and representation in ELT coursebooks, there remains limited understanding of how attitudinal meanings systematically shape stance and intercultural values in global EFL publishing. Previous research has rarely traced how linguistic patterns of Affect, Judgement, and Appreciation contribute to constructing solidarity, empathy, and global perspective in textbook discourse. Addressing this gap, the present study applies Appraisal-based discourse analysis to the *Life – Upper Intermediate* textbook, aiming to uncover how evaluative meanings function as semiotic resources for intercultural communication and stance construction in international higher-education contexts.

III. METHODOLOGY

A. Research Questions

Drawing on Appraisal-based discourse analysis, this study examines the linguistic realization of Affect, Judgement, and Appreciation in reading passages from *Life – Upper Intermediate* and addresses two research questions:

1. How are attitudinal resources (Affect, Judgement, and Appreciation) distributed and linguistically realized in the

textbook's main reading passages?

2. How do these evaluative patterns construct stance and intercultural values, and what discursive and pedagogical implications do they suggest for higher-education contexts where global materials are used?

B. Corpus

The corpus consists of the main reading passages in *Life – Upper Intermediate* (Dummett et al., 2021), a coursebook widely adopted in university programs and designed for learners at the upper-intermediate (B1+ to B2) level. To ensure comparability, only extended reading texts were included; supplementary exercises, captions, and isolated task prompts were excluded. In total, 12 units were analyzed, amounting to approximately 6,800 running words.

To provide a clearer picture of the data, Table 1 summarizes the twelve reading passages included in the corpus, highlighting their thematic scope and intercultural orientation. The selection covers a wide range of global issues – education, technology, art, psychology, environmental ethics, leadership, and innovation – thereby illustrating the pedagogical philosophy of the *Life* series: connecting language learning with knowledge of the world and shared human values.

TABLE 1
OVERVIEW OF READING PASSAGES IN *LIFE – UPPER INTERMEDIATE*

Unit	Title of reading passage	Content focus
1	<i>Bloodlines</i>	Explores migration, ancestry, and cultural identity; highlights how heritage shapes belonging and personal values.
2	<i>Once Upon a Time</i>	Examines storytelling traditions through the Brothers Grimm tales; reveals moral and cultural lessons in folk narratives.
3	<i>Appropriate Technology</i>	Discusses sustainable innovation and context-sensitive design; shows how simple technologies improve lives in developing communities.
4	<i>Why do we need music?</i>	Investigates music as an emotional and creative force; connects musical experience to empathy and cultural expression.
5	<i>The teenager mind</i>	Explores adolescent brain development and risk-taking; explains how neurological and emotional factors shape teenage behaviour and identity formation.
6	<i>Unusual places to stay</i>	Describes alternative, experiential forms of travel that promote empathy, authenticity, and intercultural understanding beyond conventional tourism.
7	<i>Cultural Conventions</i>	Analyses intercultural communication and social etiquette; emphasizes respect and adaptability across cultures.
8	<i>Saving Madagascar</i>	Profiles environmental activism and leadership; highlights conservation and ethical responsibility as global values.
9	<i>From Hero to Zero</i>	Examines media portrayal of heroism; invites critical reflection on fame, ethics, and public perception.
10	<i>The King Herself</i>	Narrates the life of Hatshepsut, the female pharaoh; illustrates women's agency and historical significance.
11	<i>Who's a Clever Bird, then?</i>	Explores scientific research on animal intelligence; links cognition, empathy, and human-animal interaction.
12	<i>Start-up</i>	Focuses on entrepreneurship and innovation; celebrates creativity, resilience, and forward-thinking as global ideals.

This textbook was selected for three main reasons. First, as an internationally distributed EFL series, *Life – Upper Intermediate* represents a prototypical example of global teaching materials designed for culturally diverse learners. Second, it is widely adopted in Asian tertiary education across various English-major and general English programmes, serving not only as a linguistic resource but also as a vehicle for cultural learning and critical reflection. Third, its reading texts, developed from authentic *National Geographic* content, combine factual reporting with attractive storytelling, providing a fertile basis for examining how evaluative language constructs stance and aligns learners with particular social values and worldviews.

C. Instruments

The primary analytical instrument was an Appraisal-based coding framework grounded in the taxonomy proposed in *The Language of Evaluation* (Martin & White, 2005). A coding sheet was used to record each identified attitudinal instance, including: (i) the evaluative expression, (ii) subsystem and subtype (Affect, Judgement, Appreciation), (iii) polarity, and (iv) the full sentence/clausal context for reference. When relevant, Graduation resources modifying Attitude were also noted.

D. Data Collection Procedures

This study adopted a qualitative discourse-analytical approach combined with basic descriptive quantification. Qualitative analysis enabled close reading of evaluative patterns in their immediate textual context, while frequency counts were incorporated to show overall tendencies across the corpus. This mixed orientation follows established practice in Appraisal-based educational research (Hood, 2010), where interpretive precision is complemented by numerical transparency.

Manual coding was deliberately chosen for its suitability to context-sensitive interpretation of evaluative meaning, which cannot be reliably automated. The procedures were as follows. First, attitudinal resources were identified at the lexical level and interpreted in their immediate clause or sentence context to ensure accurate classification. Second, each instance was entered into the coding sheet with the relevant coding fields. Third, coding decisions were verified

and refined through iterative checking against the Appraisal taxonomy to ensure consistency. The entire corpus was re-examined in multiple rounds, allowing minor adjustments and confirmation of intra-coder reliability. All data were documented transparently to allow replication and further comparative research.

IV. RESULTS AND DISCUSSION

A. Overview of Evaluative Tendencies

The analysis yielded several noteworthy tendencies. Across the corpus, a total of 358 attitudinal instances were identified, confirming that evaluative language permeates the textbook discourse. Among the three subsystems, Appreciation emerged as the most frequent category, accounting for 47.21% of all occurrences, followed by Judgement (34.92%) and Affect (17.97%) (see Table 2). This means that expressions of value occur nearly one and a half times more often than moral or emotional expressions. The predominance of Appreciation demonstrates that the textbook primarily evaluates phenomena, processes, and cultural themes rather than individuals' emotions or behaviours.

TABLE 2
DISTRIBUTION AND POLARITY OF ATTITUDE RESOURCES IN *LIFE – UPPER INTERMEDIATE*

Attitude Category	Frequency	% of total	Positive (%)	Negative (%)
Appreciation	169	47.21	79.40	20.60
Judgement	125	34.92	66.70	33.30
Affect	64	17.97	72.92	27.08
Total	358	100%	—	—

The data further reveal a strong positive orientation across all subsystems. Approximately three-quarters of all attitudinal tokens are positive in polarity, constructing a discourse of affirmation, motivation, and encouragement. For example, the statement “Getting a good education is **a must**” (Dummett et al., 2021, p. 15) assigns high social value to education and self-development. The expression “is a must” inscribes positive Appreciation (+ Valuation) by attributing necessity and high worth to the act of pursuing education. Rather than describing emotion, the clause constructs education as a socially endorsed value and a prerequisite for success. The choice of the modalized noun phrase *a must* maximizes intensity, construing education as a socially endorsed necessity and attributing high value to its pursuit. Through such formulations, the textbook promotes a stance that celebrates self-improvement and aligns learners with values of diligence and achievement.

In comparative perspective, these findings extend prior Appraisal-based research. Earlier studies of educational and textbook discourse (Hood, 2010; Geng & Wharton, 2016; Nguyen et al., 2025) revealed a stronger focus on Judgement, particularly moral or civic evaluation. In contrast, *Life – Upper Intermediate* privileges Appreciation, suggesting a shift from moral instruction to cultural and humanistic affirmation. This pattern reflects a broader transformation in global ELT publishing, where pedagogical discourse increasingly promotes intercultural awareness, social harmony, and positive engagement rather than prescriptive morality.

B. Appreciation: The Core of Evaluation

Appreciation constitutes the core of evaluation in the textbook, serving as the dominant means through which value and worth are construed. Within Appreciation, three subcategories were identified – Reaction, Composition, and Valuation – distributed by polarity in Table 3.

TABLE 3
APPRECIATION SUBCATEGORIES BY POLARITY (PER 1,000 WORDS)

Subcategory	Polarity	Freq./1000 words
Composition	Positive	0.74
	Negative	0.44
	Total	1.19
Reaction	Positive	2.07
	Negative	0.59
	Total	2.67
Valuation	Positive	16.00
	Negative	3.85
	Neutral (±)	1.19
	Total	21.04

Among these, Valuation is by far the most salient subtype, accounting for over three-quarters of all Appreciation tokens. It systematically foregrounds the value, worth, and social significance of phenomena, especially those related to education, technological innovation, cultural heritage, and social development. An illustrative example is “He wanted **affordable** technology that would lead to **greater social equality**.” [+ Valuation] (Dummett et al., 2021, p. 39).

The evaluative meanings in this clause centre on the expressions “affordable” and “greater social equality”, both of which realise (+ Valuation). The adjective “affordable” attributes worth to technology through its accessibility and usefulness rather than prestige or sophistication, defining value in practical and humanistic terms. Meanwhile, the noun

phrase “greater social equality” projects a moral dimension of value, presenting technology as a means of achieving fairness and inclusion. This configuration encapsulates the text’s humanistic stance that “appropriate technology” should empower communities and promote equitable development rather than privilege technological advancement for its own sake.

Another instance of Valuation appears in the sentence “There are hot dog and hamburger vans selling ‘**cheap eats**’ along the roadside” (Dummett et al., 2021, p. 147). The phrase “cheap eats” conveys negative Valuation, suggesting low price coupled with inferior quality. This negative appraisal is immediately contrasted with the subsequent depiction of Kogi BBQ, which is praised for offering “better-quality food at a reasonable price.” By placing these expressions side by side, the text uses the negative example as a rhetorical foil that amplifies the positive attributes of Kogi BBQ. The contrast not only highlights the value of high-quality food but also foregrounds the diversity and richness of contemporary culinary culture. Such juxtaposition demonstrates how evaluative meaning operates dialogically, constructing cultural preferences and positioning readers to appreciate particular social values.

Several other expressions in the textbook likewise realize positive Valuation, reinforcing its broadly humanistic and optimistic stance. In Unit 1, for instance, the description of the USA as “well known for being a **melting pot** of different ethnic groups and cultures” (Dummett et al., 2021, p. 15) constructs cultural diversity and social cohesion as admirable national attributes. In Unit 2, the assertion that the Grimm collection “should be used as ‘a **manual of manners**’” (Dummett et al., 2021, p. 27) elevates the tales’ moral and didactic purpose, framing them as culturally enduring and socially instructive. Similarly, in Unit 4, the claim that music “can **make** a verbal message **stronger**” (Dummett et al., 2021, p. 51) highlights its communicative and affective power. Together, these expressions foreground the value of cultural richness, ethical guidance, and artistic influence, illustrating how positive Valuation is deployed to endorse socially desirable qualities throughout the textbook.

The predominance of Valuation thus demonstrates the textbook’s orientation toward promoting optimism, progress, and social contribution. Rather than encouraging critique or detachment, the discourse cultivates an affirmative worldview that celebrates what is admirable or worth emulating. While this tendency aligns with patterns observed in many ELT and educational materials, where positive evaluative meanings help construct culturally affirmative narratives (e.g., Gray, 2010; Tomlinson, 2012; Nguyen et al., 2018; Nguyen et al., 2025; Nguyen, 2023), it contrasts with findings from other genres. In history textbooks, for example, Judgement frequently dominates, particularly in the appraisal of human actions, responsibilities, and moral character (Coffin, 2006; Myskow, 2017; Nguyen et al., 2025). Similarly, political and argumentative discourse often foregrounds critical, adversarial, or morally charged evaluations (Ross, 2016; Termjai, 2025). The prominence of Valuation in the present corpus therefore highlights a comparatively humanistic and intercultural evaluative orientation, distinguishing ELT textbook discourse from other textbook genres and political discourse, where Judgement and more overtly moral or adversarial evaluation tend to be more central.

Reaction and Composition occurred far less frequently but nonetheless contributed subtle layers to the evaluative texture of the discourse. Reaction, expressed through affective adjectives such as *amazing*, *impressive*, *exciting*, and *rousing*, evokes emotional resonance, while Composition, realized through references to clarity or balance (e.g., *simple*, *work simply*, *understandable*), assesses the structural and aesthetic qualities of phenomena. Across all subcategories, positive polarity overwhelmingly dominated, producing a discourse of affirmation and admiration that privileges human progress, creativity, and cultural enrichment. This tendency aligns with broader findings in research on ELT and educational materials, where positive evaluative meanings often reinforce culturally affirmative narratives and promote humanistic, optimistic values.

C. Judgement and Affect: Supporting Interpersonal and Emotional Functions

Table 4 summarizes the subcategories of Judgement and Affect together with their polarity values normalized per 1,000 words.

TABLE 4
JUDGEMENT AND AFFECT SUBCATEGORIES BY POLARITY (PER 1,000 WORDS)

Subsystem	Subcategory	Positive	Negative	Neutral (±)	Total
Judgement	Capacity	4.59	1.04	—	5.63
	Normality	1.19	1.33	—	2.52
	Propriety	3.70	3.26	0.15	7.11
	Tenacity	2.37	—	—	2.37
	Veracity	0.30	0.44	—	0.74
Affect	Happiness	1.19	0.74	—	1.93
	Inclination	0.74	0.15	—	0.89
	Satisfaction	3.70	0.15	—	3.85
	Security	0.74	1.33	0.15	2.22

While Appreciation predominated, Judgement and Affect served complementary interpersonal and emotional functions within the evaluative configuration of the textbook. Judgement emerged as the second-largest subsystem, with most instances clustering around Propriety and Capacity. Positive Judgement frequently highlighted ethical responsibility and competence, as in the sentence “He has **persuaded** the locals **to stop cutting down trees**” (Dummett et al., 2021, p. 39). The phrase “persuaded to stop cutting down trees” encodes positive Propriety, foregrounding

environmentally responsible behaviour and communal benefit. In contrast, negative Propriety targeted negligence or harmful practices, exemplified by the description of a "new leader with **little interest in protecting the environment**" (Dummett et al., 2021, p. 39). Here, "little interest in protecting the environment" expresses negative Propriety, condemning indifference to ecological issues.

Rather than foregrounding harsh critique, the textbook relied mainly on positive Propriety and Capacity to promote responsibility, competence, and environmentally ethical behaviour. This pattern aligns with observations that global ELT materials frequently embed "soft" moral narratives and avoid ideological confrontation in order to remain culturally acceptable across diverse markets (Gray, 2010; Tomlinson, 2012). At the same time, the relatively restrained use of negative Propriety underscores an orientation toward encouragement rather than blame, reflecting the genre's pedagogical aim of fostering global citizenship and cooperative values.

Affect, although the least frequent subsystem, played a strategic role in humanizing narratives and fostering empathy. The most salient subtype was positive Satisfaction, as seen in the portrayal of multicultural communities where "second-generation Puerto Ricans live alongside third-generation Greeks and first-generation Koreans, all united by a **common feeling of pride** in their American identity" (Dummett et al., 2021, p. 15). By depicting second-, third-, and first-generation immigrant groups living alongside one another, the text frames collective satisfaction as arising from a shared sense of belonging that transcends generational and cultural differences. Negative Affect appeared only sporadically, serving pragmatic guidance rather than emotional intensity. For example, "Bear in mind that if you stay during the week, you might be **disappointed**" [– Satisfaction] (Dummett et al., 2021, p. 75) uses the adjective "disappointed" to signal unmet expectations and to guide readers toward realistic anticipation during their vacation stay.

Other subtypes occurred with lower frequency but contributed diversity. Happiness included positive items such as *emotional reunion*, *loved*, *happily ever after*, and *cheerful optimism*, alongside negatives like *didn't like*, *betrayed* and *desperate*, and *depressed*, indicating a balanced portrayal of emotion. Security contrasted stability (e.g., *comforted*, *calming*) with vulnerability (e.g., *caught in a trap*, *alarmed*, *at risk*). Satisfaction was mainly positive (e.g., *pride*, *proud*, *celebrate*, *feels like a treat*), with occasional negatives (e.g., *disappointed*). Inclination was mostly positive (e.g., *ambition*, *prefer the company*), with infrequent negatives (e.g., *refuse*).

The distribution of Judgement and Affect indicates a complementary interpersonal orientation that supports the strong presence of Appreciation in the corpus. Judgement functioned as a resource for shaping moral positioning in a relatively soft and constructive manner, while Affect contributes a restrained emotional dimension, primarily through positive Satisfaction, without introducing strong emotional intensity. The scarcity of negative Affect strengthens the textbook's affirmative tone and avoids emotional dissonance. These patterns resemble tendencies in other educational contexts where emotional alignment is used to scaffold learner motivation and promote inclusive identities (Nguyen, 2023).

V. CONCLUSION

This study examined the distribution and realization of attitudinal resources in *Life – Upper Intermediate*, an internationally distributed EFL textbook widely used in higher education. The findings reveal a clearly affirmative evaluative profile across the corpus. Among the three Attitude subsystems, Appreciation – particularly Valuation – emerged as the dominant resource, while Judgement played a secondary but systematic role and Affect occurred less frequently and in a more controlled manner. The predominance of Appreciation together with a strong positive polarity indicates an evaluative configuration that directs attention toward positively valued phenomena and cultural practices, shaping an affirmative stance toward the world. These findings carry important theoretical and discursive implications.

A. Theoretical Implications

This analysis contributes to Appraisal-based research by demonstrating how attitudinal meanings are systematically embedded in the discourse of EFL textbooks, a genre of growing global significance in tertiary education. The dominance of Appreciation corroborates previous Appraisal analyses of academic and instructional discourse, where evaluation typically targets processes and phenomena rather than persons. However, in contrast with national or locally authored textbooks that rely heavily on Judgement to transmit moral or civic norms, *Life – Upper Intermediate* privileges Appreciation to celebrate learning, innovation, and cultural achievement. This pattern suggests a broader tendency toward epistemic celebration rather than overt moral prescription in contemporary global ELT discourse.

From a stance perspective, the study enriches existing research by situating stance within the discourse-semantic system of Attitude. The cumulative pattern of evaluation suggests that stance is constructed primarily through appraisal attached to phenomena and practices, rather than through explicit authorial engagement. In this way, the study bridges Appraisal Theory and stance research, extending stance analysis from lexical and syntactic realizations to a systemic account of meaning, and highlighting stance as a socially embedded and ideologically charged phenomenon in educational discourse.

B. Discursive and Intercultural Implications

The evaluative configuration identified in *Life – Upper Intermediate* also reveals how stance and ideology are jointly constructed at the discursive level. The textbook's evaluative language operates as an affiliative framework, gradually

aligning readers with an affirmative worldview rather than engaging in argumentation. Recurrent positive Appreciation, particularly the Valuation of education, science, and cultural achievement, invites learners to participate affectively in narratives of progress and cooperation. Judgement resources further reinforce an image of humanity defined by competence, responsibility, and goodwill, while the restrained use of Affect maintains a controlled emotional tone.

These patterns realize what may be described as an affirmative intercultural stance, one that prioritizes harmony, cooperation, and shared progress while backgrounding conflict or contestation. Intercultural understanding is constructed not through the negotiation of difference, but through the normalization of shared virtues and universally desirable goals. Even negative evaluations – such as criticism of pollution, ignorance, or corruption – function to reaffirm improvement and inclusion rather than to polarize perspectives. This confirms that evaluation operates as a macro-level stance resource, sustaining interpersonal alignment through the thematic and lexical architecture of the text rather than through explicit authorial presence.

Pedagogically, recognizing such evaluative framing is crucial for fostering critical literacy. When learners are guided to examine how Appreciation and Judgement encode cultural assumptions and value hierarchies, they become better equipped to recognize how positivity functions to universalize certain perspectives while marginalizing others. For materials designers and teachers, these insights suggest that explicit reflection on evaluative language can deepen intercultural competence, transforming affirmation into critical awareness rather than passive alignment.

C. Recommendations for Future Research

Despite its contributions, this study has several limitations that point to directions for future research. First, the analysis focused on a single textbook and was restricted to the Attitude subsystem, leaving Engagement and Graduation unexplored. Future studies could extend the analysis to these subsystems to provide a more comprehensive account of evaluative meaning-making in instructional discourse. In particular, examining Engagement resources would help clarify how dialogic space, alignment, and alternative viewpoints are managed in international EFL textbooks, especially in relation to the presentation of cultural and social issues. Second, comparative research across textbooks of different proficiency levels or publishing traditions would help determine whether the affirmative evaluative profile observed here is characteristic of global EFL materials more broadly.

REFERENCES

- [1] Channell, J. (2000). Corpus-based analysis of evaluative lexis. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 38–55). Oxford University Press.
- [2] Chou, M. H. (2022). Using literature circles to teach graded readers in English: An investigation into reading performance and strategy use. *Innovation in Language Learning and Teaching*, 16(2), 144–163. <https://doi.org/10.1080/17501229.2021.1885412>
- [3] Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. Continuum.
- [4] Dontcheva-Navratilova, O. (2009). Epistemic modality in Czech academic writing in English. *Discourse and Interaction*, 2(2), 7–24.
- [5] Dummett, P., Hughes, J., & Stephenson, H. (2021). *Life: Upper intermediate* (2nd ed.). National Geographic Learning/Cengage Learning.
- [6] Ebrahimi, S. F., & Mohsenzadeh, A. (2022). Functional analysis of stance nouns in English Language Teaching research articles. *Southern African Linguistics and Applied Language Studies*, 40(4), 430–440. <https://doi.org/10.2989/16073614.2022.2075903>
- [7] Geng, Y., & Wharton, S. (2016). Evaluative language in doctoral theses: Similarities and differences between L1 Chinese and L1 English writers. *Journal of English for Academic Purposes*, 22(1), 80–91. <https://doi.org/10.1016/j.jeap.2016.01.001>
- [8] Gray, J. (2010). *The construction of English: Culture, consumerism and promotion in the ELT global coursebook*. Palgrave Macmillan.
- [9] Hashim, S. S. (2022). Critical discourse analysis of the attitudinal resources in modern English poetry: The appraisal system. *International Journal of Linguistics*, 14(2), 163–181.
- [10] Hood, S. (2010). *Appraising research: Evaluation in academic writing*. Palgrave Macmillan.
- [11] Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
- [12] Hyland, K. (2017). Metadiscourse: What is it and where is it going? *Journal of Pragmatics*, 113, 16–29. <https://doi.org/10.1016/j.pragma.2017.03.007>
- [13] Hyland, K., & Jiang, F. K. (2016). “We must conclude that...”: A diachronic study of academic engagement. *Journal of English for Academic Purposes*, 24(1), 29–42. <https://doi.org/10.1016/j.jeap.2016.09.003>
- [14] Khamkhien, A. (2014). Linguistic features of evaluative stance: Findings from research article discussions. *Indonesian Journal of Applied Linguistics*, 4(1), 54–69. <https://doi.org/10.17509/ijal.v4i1.600>
- [15] Kim, C. (2025). Evaluative stance in L1 and L2 argumentative essays: A corpus-based comparison. *Forum for Linguistic Studies*, 7(4), 998–1011.
- [16] Martin, J. R., & Veel, R. (Eds.). (1998). *Reading science: Critical and functional perspectives on discourses of science*. Routledge.
- [17] Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- [18] Muti'ah, & Azizah, V. M. (2024). Investigating students' attitudes on oral corrective feedback: Small-scale survey in academic speaking class. *International Journal of English Linguistics, Literature, and Education (IJELLE)*, 6(1), 26–37.
- [19] Myskow, G. (2017). Appraising history: Stance and engagement in Japanese high school history textbooks. *Linguistics and Education*, 38, 1–10.
- [20] Nguyen, B. H., Vu, H. T., Bui, T. T. T., Vu, T. T. T., & Dang, H. T. K. (2025). Evaluative language: Attitudinal resources in

- Vietnamese history textbooks. *Theory and Practice in Language Studies*, 15(2), 656–666. <https://doi.org/10.17507/tpls.1502.35>
- [21] Nguyen, T. H. L., Pham, L. T., Nguyen, T. T., & Ngo, B. T. (2018). Evaluative language in Singaporean and Vietnamese primary school textbooks (a preliminary study on Grade 2). *HNUe Journal of Science: Educational Sciences*, 67(2), 22–28.
- [22] Nguyen, T. M. L. (2023). *The language of evaluation in English textbooks at the secondary level in Vietnam from the viewpoint of systemic functional grammar* [Doctoral dissertation]. University of Education, The University of Danang.
- [23] Ross, A. S. (2016). The construction of stance in the 2012 Australian prime ministerial debates. *Journal of Language and Politics*, 15(4), 488–510.
- [24] Sahragard, R., & Meihami, H. (2016). A selective study of the appraisal system in cosmetic advertisements: A systemic functional perspective. *International Journal of Applied Linguistics and English Literature*, 5(4), 45–52.
- [25] Siromaa, M., & Rauniomaa, M. (2021). Stance and evaluation. In M. Haugh, D. Z. Kádár, & M. Terkourafi (Eds.), *The Cambridge handbook of sociopragmatics* (pp. 95–116). Cambridge University Press.
- [26] Termjai, M. (2025). A study of vague language in argumentative essays by Thai EFL undergraduate students. *3L: Language, Linguistics, Literature*, 31(1), 269–284.
- [27] Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 1–27). Oxford University Press.
- [28] Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, 45(2), 143–179.
- [29] Vo, L. H., & Le, L. V. (2025). Exploring EFL writers' authorial voice in academic discussions: A graduation-based appraisal analysis. *Forum for Linguistic Studies*, 7(5), 932–946.
- [30] White, P. R. R. (1998). *Telling media tales: The news story as rhetoric* [Doctoral dissertation]. University of Sydney.
- [31] White, P. R. R. (2006). Evaluative semantics and ideological positioning in journalistic discourse. In R. Miller & M. Pano Alamán (Eds.), *Mediating ideology in text and image* (pp. 37–67). Peter Lang.
- [32] White, P. R. R. (2016). Evaluative contents in verbal communication. In A. Rocci & L. de Saussure (Eds.), *Verbal communication* (Vol. 3, Handbooks of Communication Science, pp. 77–96). De Gruyter Mouton.



Ngoc T. B. Han was born in Vietnam in 1988. She received her PhD in Linguistics from the Vietnam Academy of Social Sciences, Hanoi, Vietnam, in 2025.

She is currently a Lecturer at Thuongmai University, Hanoi, Vietnam, where she teaches linguistics and English for Academic Purposes. Her research focuses on discourse studies, including point of view, narration, and evaluation in educational and literary discourse.

Dr. Han is a member of the academic staff at Thuongmai University and has contributed to various institutional projects on language education and curriculum development. For more information, contact: ngoc.htb@tmu.edu.vn

The Impact of Cultural Nuances in the Translation of Indonesian Government Social Media Content on the Digital Literacy of Young Generations

Burhanuddin Arafah

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Herawaty Abbas

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Muhammad Hasyim

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Muhammad Ridha Anugrah Latief

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Taqdir

Faculty of Cultural Sciences, Hasanuddin University, Makassar, Indonesia

Abstract—This study examines the influence of cultural nuances in the translation of Indonesian government social media content on the digital literacy of young generations. Using a quantitative approach, data were collected from 243 respondents aged 17–22 years through an online questionnaire. The analysis covered cultural nuances (X1), translation strategies (STR) as mediator, machine translation (MT) and socio-political context (CTX) as moderators, message comprehension (X2), and digital literacy (Y). The results indicate that cultural nuances positively influence youth comprehension, while translation strategies play a limited mediating role. Machine translation and socio-political context significantly moderate how messages are interpreted. Furthermore, digital literacy is generally high in terms of ethics and safety but remains low in creation and participation. These findings highlight the need for more contextual translation strategies and greater efforts to empower youth as active digital participants or contributors.

Index Terms—translation, cultural nuances, social media, government communication, digital literacy

I. INTRODUCTION

The rapid expansion of digital technologies has transformed how governments communicate with citizens. Public institutions increasingly rely on social media to disseminate information, promote transparency, and engage with diverse audiences (Arafah & Hasyim, 2019; Hasjim et al., 2026). In Indonesia, platforms such as Instagram, X (formerly Twitter), TikTok, and YouTube have become central channels for policy dissemination, outreach, and civic participation. Within this multilingual and multicultural context, where linguistic diversity and cultural plurality are integral to national identity (Kaharuddin et al., 2024), the translation of government messages is not merely a technical exercise but a sociocultural process (Arafah et al., 2026). Numerous cultural studies have shown that local traditions, symbolic beliefs, and community-based practices strongly shape communication patterns and identity construction in Indonesian society (Arafah et al., 2022; Arifin et al., 2022; Effendy et al., 2025; Takwa et al., 2024a; Takwa et al., 2025). Language and texts do not merely communicate information, but also reflect social realities, ideological tensions, and unequal relations within society (Mutmainnah et al., 2022). Texts and communicative forms are often shaped by wider social structures, including hierarchy, class positioning, and culturally sanctioned roles (Rano et al., 2026). How these messages are translated determines how citizens interpret, respond to, and trust official communications, particularly among younger generations, who constitute the most active users of digital media (Arafah et al., 2025a; Hasyim & Arafah, 2023a).

Environment changes when the concept of cultural heritage is fading due to changes in lifestyle, modernization, and technological developments (Manugeren et al., 2023). The shift from what once traditional into a modern one is possible to perform both good and bad impacts (Suhadi et al., 2022). Effective translation in government communication is crucial for bridging the gap between official discourse and the cultural realities of its audience (Hasyim & Arafah, 2023b). Cultural nuances encompassing expressions, idioms, values, and historical references carry implicit meanings that influence the emotional and cognitive engagement of the reader (Arafah et al., 2023a; Arafah et al., 2025b). When these nuances are inadequately translated, messages risk appearing inauthentic, being misunderstood, or failing to engage audiences. The digital sphere amplifies these challenges, as communication on social media occurs rapidly, interactively,

and across linguistic and cultural boundaries. Moreover, the increasing reliance on machine translation in public communication raises concerns about the effectiveness of cultural meaning preservation, potentially compromising the authenticity and inclusiveness of government messages. Similar observations have been made in cross-cultural linguistic research, which highlights the importance of contextual interpretation and the negotiation of meaning in maintaining cultural coherence (Mokoginta & Arafah, 2022).

Beyond linguistic accuracy, cultural translation reflects how collective memory, moral values, and social worldviews are encoded in language. Research by Radjaban et al. (2025) reveals that historical background of Merap Language of Dayak community is used as a benchmark to determine a language using language mapping analysis. More by Halil et al (2024), it is stated that North Kolaka has districts that are rich culture and language. Studies in ethnolinguistics and cultural semiotics show that meaning in Indonesian local communities—such as the Tolaki Mekongga and Biak tribes—is inseparable from symbolic rituals, ecological wisdom, and moral communication (Mofu et al., 2024; Takwa et al., 2024). Value system, including rituals, symbols, and traditions is known to be a guideline of how a community acts and behaves to make their life better (Takwa et al., 2022a; Takwa et al., 2026; Abbas et al., 2024). These cultural models illustrate how social messages are transmitted through contextualized linguistic codes and nonverbal symbolism, reinforcing local identity while shaping intercultural understanding. In public communication, similar semiotic dynamics emerge when governments attempt to localize national values such as gotong royong and Pancasila through translated messages. Governments should know that general people as their audience will interpret a term based on the meaning they understand (Hasjim et al., 2020). As cultural symbols move across linguistic boundaries, they undergo reinterpretation that either strengthens or weakens audience trust and engagement. Thus, government translation can be seen as an act of cultural negotiation—mediating between institutional authority and the lived experience of its citizens (Takwa et al., 2022b).

In a world of communication, people are expected to create social interaction effectively despite all the implicit meanings the speaker assumes (Yulianti et al., 2022; Kuswanti et al., 2023; Asri et al., 2023). At the same time, digital communication has reshaped how meaning is created, shared, and contested. Scholars note that in the new media environment, multimodal signs—text, visuals, and emotive symbols—function collaboratively to construct meaning beyond linguistic boundaries (Hasyim & Arafah, 2023b; Arafah & Hasyim, 2023a). Semiotic multimodality plays a critical role in how audiences interpret official messages online, as emojis, imagery, and digital tone affect perceived sincerity and clarity. Research in digital literacy and media semiotics also suggests that communicative meaning on social platforms is increasingly shaped by audience participation, visual culture, and algorithmic mediation (Arafah et al., 2025c). Research by Iksora et al. (2022) shows that in the use of language, an error or mistake made by either human or program is unavoidable. Using metaphors, for instance, in order to convey ideas might just end up with misunderstanding (Baa et al., 2023). Consequently, translation in digital government communication must consider these multimodal dimensions to ensure that the emotional, cultural, and ethical layers of meaning are preserved. Learning a new language is not an easy job comparing to people's first language (Kaharuddin et al., 2023). The complexity of understanding translation from another language is an obstacle (Arnawa & Arafah, 2023). Mastering a foreign language remains a significant challenge (Kaharuddin et al., 2025). Failure to do so can result in the erosion of cultural resonance and the weakening of civic trust in digital public discourse.

Simultaneously, digital literacy has emerged as an essential skill for youth to access, evaluate, and interact with online information (Usman et al., 2024). According to UNESCO, digital literacy encompasses the ability to access, evaluate, create, and ethically use information in digital environments. In this context, government social media communication not only functions as a channel of information delivery but also serves as a medium for fostering citizens' critical digital engagement (CDE). A growing body of research emphasizes that integrating local wisdom and cultural perspectives into education and communication practices strengthens intercultural awareness and digital competence among young people (Misnah et al., 2024). Moreover, the modernization era also gives an impact to educational sphere to be able to access online effectively (Arafah et al., 2023b; Sunardi et al., 2018). Integrating cultural awareness to the teaching materials could be easily done since it is presented in many forms, literary work for instance (Yudith et al., 2023). Authors occasionally depict social background and phenomena around them, creating a close relationship between literature and social reality (Sunyoto et al., 2022; Arafah et al., 2023). In Indonesia, however, youth digital literacy levels vary widely, with gaps in evaluative and creative competencies despite high connectivity rates. Recent findings also show that Generation Z netizens demonstrate strong technological adaptability but tend to lack a deeper reflective and ethical understanding of digital content, indicating an uneven development of digital cultural literacy (Arafah et al., 2023c; Arafah & Hasyim, 2023b, 2023a).

Previous studies have mostly examined translation in government communication from linguistic or pragmatic perspectives, focusing on accuracy, equivalence, and politeness strategies (e.g., Venuti, 1995; Schäffner, 2012). Few studies address how translation interacts with cultural values and affects audience perception, especially in digital contexts. Recent research on digital government communication emphasizes that meaning-making in online spaces is co-constructed between institutions and citizens, mediated by sociocultural interpretation (Marwaha & Kvedar, 2021). Related studies on digital media and multimodal communication show that online meaning construction relies on semiotic, visual, and emotional cues shaping users' interpretive engagement (Arafah & Hasyim, 2023b; Hasyim & Arafah, 2023b). Research on digital literacy among youth (Arafah & Hasyim, 2023b) reveals that although young users are digitally active,

they often lack the critical and evaluative skills to interpret messages with ideological and cultural dimensions. This suggests a need to investigate how translation quality and cultural sensitivity shape comprehension and trust in official information sources.

Despite interest, little is known about how cultural nuances in translation affect digital literacy in multilingual societies like Indonesia. Most studies treat translation and digital literacy separately, ignoring how language, culture, and technology influence digital participation. Increasing machine translation use for official communication complicates preserving cultural meaning. This study integrates translation studies, intercultural communication, and media literacy research, contributing to discourse on culturally responsive translation enhancing young people's critical engagement with government social media. It examines relationships among cultural nuances, comprehension, and digital literacy, offering insights for improving public digital communication. This study investigates (1) how cultural nuances affect comprehension, (2) if translation strategies mediate this, (3) if machine translation and socio-political context moderate comprehension, and (4) how these factors shape youth digital literacy.

II. LITERATURE REVIEW

A. *Translation and Cultural Nuances*

Translation is a cultural act (Jamiluddin et al., 2026). Venuti (1995) argues translation is not neutral linguistic substitution but a cultural negotiation reflecting power, ideology, and identity. Translators mediate between languages and cultures through domestication—adapting expressions to target-culture norms—and foreignization—preserving source-culture specificity. Domestication makes messages accessible but risks erasing cultural distinctiveness; foreignization maintains authenticity but may hinder comprehension for readers unfamiliar with the source culture (Venuti, 1998; House, 2015).

Recent studies on cultural representation emphasize that translation and educational media often reproduce cultural hierarchies rather than challenge them (Taqdir et al., 2026). Visual and textual representations in foreign language textbooks, for example, reinforce cultural and gender stereotypes by privileging dominant cultures while marginalizing local ones (Taqdir, 2025). This highlights that translation and representation are not neutral acts but embedded in ideological systems shaping how meaning and identity are negotiated across languages.

In Indonesia, studies of linguistic and cultural expression in education and communication show that negotiation and meaning-making are rooted in cultural values like harmony, politeness, and collective respect. Mokoginto and Arafah (2022) show that Indonesian communication patterns are shaped by cultural schemas emphasizing indirectness and social harmony. These findings align with translation studies stressing how cultural knowledge informs lexical choice and pragmatic tone in intercultural communication.

Nida's (1964) dynamic equivalence theory and Hall's (1976) high- and low-context communication highlight that meaning relies on shared cultural background and implicit values. In multilingual contexts like Indonesia, cultural nuances—idioms, values, and symbols in local discourse—determine how messages are emotionally and cognitively received (Takwa et al., 2026). In multilingual and multicultural settings, meaning is influenced by lexical choice, socially constructed expectations, gendered norms, and culturally embedded values (Yudith et al., 2024). Similarly, Arafah et al. (2023b) argue that effective communication among speakers of different linguistic and cultural backgrounds requires mutual cultural understanding and empathy, shaping how messages are interpreted in multilingual societies. Failure to convey these nuances may result in misinterpretation or perceived insincerity, especially when messages are translated for public dissemination. In government communication, this may lead to a decline in public trust or engagement, as language becomes a marker of inclusivity and representation (Schäffner, 2012; Hatim & Mason, 2014). Thus, culturally informed translation strategies are essential for preserving message integrity and ensuring communicative resonance across linguistic and cultural boundaries.

B. *Government Communication in Social Media*

Governments worldwide increasingly use social media to interact with citizens, share information, and promote transparency (Bonsón et al., 2012). The shift from broadcasting to dialogue has redefined public communication. However, effective engagement relies on technological access, linguistic clarity, cultural sensitivity, and contextual framing (Criado & Villodre, 2022). Social media creates new expectations for authenticity and immediacy, making tone, wording, and cultural relevance critical for message perception and sharing. In multilingual societies like Indonesia, government communication is tied to cultural representation. Translation must account for local expressions, cultural hierarchies, and sociopolitical nuances to maintain credibility. Overly formal, literal, or culturally detached messages, often due to machine translation, may seem impersonal or unrelatable, reducing trust and civic participation (Nittas et al., 2024). Therefore, government communicators must balance linguistic precision with cultural adaptability to sustain digital citizen engagement.

C. *Digital Literacy*

Digital literacy has shifted from technical skills to include critical, creative, and ethical engagement with online content (Ng, 2012). UNESCO defines it as the ability to access, evaluate, create, and ethically use digital information. These skills enable individuals to navigate online, assess message credibility, and contribute to digital discourse. In Indonesia, digital

literacy addresses misinformation and communication challenges on social media, where users' interpretive skills are shaped by culture and online exposure (Arafah & Hasyim, 2023b). Studies show young users have high connectivity but uneven digital literacy (Idhan et al., 2026). While proficient in consuming digital content, their evaluation, creation, and civic skills are limited. This imbalance challenges interpretation of government messages, requiring critical understanding of sociocultural and political nuances. As Yotha (2025) notes, digital literacy involves accessing information and developing reflective, participatory capacities shaping democratic engagement. Thus, fostering digital literacy among youth is crucial for building informed, critically aware digital citizens.

D. Interconnection Between Translation, Culture, and Digital Literacy

Although translation and digital literacy have been extensively studied, their combined examination is rare. The interplay between cultural translation and digital engagement is underexplored, especially in governmental communication. Cultural nuances in translated texts influence audience comprehension and evaluation, shaping digital literacy practices (Da Silva, 2014; Shifman, 2013). Inaccurate or culturally insensitive translations can distort meaning, reduce trust, and limit critical engagement with official messages. Conversely, culturally responsive translation enhances interpretive depth, inclusivity, and empowers youth in digital governance. Takwa et al. (2025) emphasized that language in ritual communication reflects cultural values like respect, ethics, and harmony, guiding interpretation and behavior. Arafah et al. (2023) illustrated that cultural imagery and symbolism in rituals function as frameworks sustaining identity and moral cohesion. These findings suggest cultural meaning in local discourse, oral or digital, shapes interpretive behavior and engagement similarly. This study is at the intersection of translation studies, intercultural communication, and media literacy. By integrating these perspectives, it examines how translation strategies and cultural nuances contribute to message comprehension and digital literacy among young Indonesians. This integration offers a comprehensive lens for understanding how linguistic, cultural, and technological factors shape digital participation in governance.

E. Research Gap

Despite increasing attention paid to digital government communication and literacy, three major gaps persist. First, most studies address translation or digital literacy separately, neglecting their interconnection within sociocultural communication processes. Second, limited research has examined these relationships empirically, particularly using quantitative models that test mediating and moderating factors such as translation strategy, machine translation, and socio-political context. Third, there is little evidence from Southeast Asia, where multilingualism and cultural diversity significantly affect message reception. Recent bibliometric evidence also shows that, while global research in language education has increasingly explored intersections among technology, culture, and identity, studies rarely extend these insights to public or governmental digital communication (Taqdir et al., 2025). This study fills these gaps by empirically investigating how cultural nuances in the translation of Indonesian government social media content influence youth digital literacy with message comprehension as a key intermediary variable.

III. METHODOLOGY

A. Research Design

This study employed a quantitative research design using a cross-sectional online survey to examine the relationships among cultural nuances, translation strategies, machine translation, socio-political context, message comprehension, and digital literacy. A quantitative approach was appropriate because the study sought to identify causal associations and test the mediating and moderating effects among constructs. A quantitative approach was appropriate because the study aimed to examine relationships among variables and to test indirect effects within a language-education context (Suparti et al., 2025). The cross-sectional design allowed data to be collected at a single point in time, reflecting participants' perceptions and experiences in the natural context of social media engagement. Descriptive and inferential statistical analyses were conducted to interpret both the general patterns and the hypothesized relationships among variables.

B. Respondents

A total of 243 respondents aged between 17 and 22 years participated through purposive sampling. The sample size was considered adequate according to statistical requirements for multiple regression analysis (Hair et al., 2019). Female respondents constituted 138 (56.79%) and male respondents 105 (43.21%). Regarding educational background, 216 participants (88.89%) had completed senior high school or its equivalent, followed by 24 (9.88%) bachelor's degree, and 3 (1.23%) diploma holders.

Participants were selected based on two main criteria: (1) active engagement with social media platforms such as Instagram, X (formerly Twitter), TikTok, or YouTube, and (2) exposure to translated government communication online. The data were collected using an online questionnaire distributed through social media and university networks. Participation was voluntary, and informed consent was obtained prior to participation. All responses were anonymized to ensure ethical compliance and data confidentiality.

C. Instrument

The questionnaire was structured into three main sections: (1) demographic information, (2) items measuring cultural nuances, translation strategies, machine translation, socio-political context, and message comprehension, and (3) digital literacy dimensions adapted from UNESCO's (2018) framework. Each construct was operationalized using multiple indicators derived from established theories (Venuti, 1995; Nida, 1964; Hall, 1976; UNESCO, 2018). All items were rated on a 5-point Likert scale ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree").

A pilot test with 30 participants was conducted to ensure clarity and content validity of the items. Cronbach's Alpha values from the pilot test indicated satisfactory internal consistency ($\alpha > 0.70$). Minor adjustments were made to item wording to improve readability and contextual relevance. The final instrument contained 33 items across six variables.

D. Variables

(a) Independent Variable (X1): Cultural nuances in translation (5 items).

1. The translation of government content retains uniquely Indonesian cultural terms (e.g., *gotong royong*, *Pancasila*). (X1_1)
2. The translation explains the meaning of cultural terms that may be unclear to non-local audiences. (X1_2)
3. The translation style feels appropriate to the Indonesian cultural context (not rigid or awkward). (X1_3)
4. The translation preserves important historical or cultural symbols/references. (X1_4)
5. I feel that the translation successfully maintains the emotional nuances of Indonesian culture (e.g., respect, togetherness). (X1_5)

(b) Mediator: Translation strategies (STR, 5 items).

1. The translation tends to make foreign terms sound more "familiar" (domestication). (STR_1)
2. The translation often retains foreign elements without significant modification (foreignization). (STR_2)
3. The translation focuses on the meaning and communicative effect (dynamic approach) rather than word-for-word structure. (STR_3)
4. The translation maintains the formal style typical of government communication. (STR_4)
5. I often find additional notes or explanations (footnotes or captions) for cultural terms in the translation. (STR_5)

(c) Moderator 1: Machine translation (MT, 5 items).

1. I believe that much of the government's translated content is produced/translated by machines (e.g., Google Translate). (MT_1)
2. Automatically labeled translations (e.g., "auto-translated") tend to miss cultural nuances. (MT_2)
3. When the translation is done by humans, I feel that cultural quality and meaning are better preserved. (MT_3)
4. I have encountered clear meaning errors caused by machine translation. (MT_4)
5. I trust translations that are edited/reviewed by humans more than purely machine-generated ones. (MT_5)

(d) Moderator 2: Socio-political context (CTX, 5 items).

1. I tend to accept government messages when the political or social context is stable. (CTX_1)
2. The sensitivity of social or political issues affects how I interpret government messages. (CTX_2)
3. My level of trust in government institutions influences how clearly I understand the message. (CTX_3)
4. Social media often displays public reactions (positive or negative) that affect my interpretation of the message. (CTX_4)
5. Issues of identity or culture (ethnicity, religion, locality) influence my reception of the government's translated messages. (CTX_5)

(e) Intervening Variable (X2): Message comprehension (5 items).

1. I find translated government messages easy to understand. (X2_1)
2. The translation helps me grasp the implied meaning (beyond the literal meaning). (X2_2)
3. I feel that the relevance of the message to my life as a young person is clear. (X2_3)
4. After reading the translation, I am able to explain the main idea of the message to others. (X2_4)
5. I tend to trust official messages when the translation feels clear and contextual. (X2_5)

(f) Dependent Variable (Y): Digital literacy (8 items).

Access & Navigation

1. I can easily find official government information on digital platforms. (Y_Akses1)
2. I know how to use platform features (search, filter) to access official content. (Y_Akses2)

Information Evaluation

1. I am able to assess the credibility of government messages I find online. (Y_Eval1)
2. I can distinguish between official information and rumors on social media. (Y_Eval2)

Creation & Participation

1. I have created or shared content in response to government messages (comments, reposts, videos). (Y_Kreasi1)
2. I have participated in online discussions related to official government content. (Y_Kreasi2)

Digital Safety & Ethics

1. I am aware of the importance of checking sources before sharing information. (Y_Etika1)
2. I am careful about privacy and personal data when interacting with government content. (Y_Etika2)

E. Data Analysis

Data analysis was performed using IBM SPSS 26 and SmartPLS 4. The analytical procedures included three stages.

First, descriptive statistics (mean, standard deviation, frequency) were computed to summarize respondent profiles and variable tendencies.

Second, reliability and validity tests were conducted, including Cronbach's Alpha, composite reliability, and average variance extracted (AVE) to ensure measurement quality.

Third, inferential analysis was performed using multiple regression and structural equation modeling (SEM-PLS) to test direct, mediating, and moderating relationships. Mediation effects were assessed using bootstrapping with 5,000 resamples, while moderation effects of machine translation (MT) and socio-political context (CTX) were examined through interaction terms.

All analyses adopted a 95% confidence interval ($\alpha = 0.05$). Model fit and multicollinearity were checked prior to interpretation.

IV. RESULTS

This section presents the main findings of the study, beginning with the respondents' social media usage patterns, followed by descriptive statistics for each variable, reliability analysis, and inferential testing of the proposed relationships.

A. The Social Media Usage

Most respondents reported spending more than four hours per day (62.14%) on social media, indicating a high degree of digital engagement among Indonesian youth. Only 0.41% of the participants reported never using social media. The most frequently used platforms were Instagram (37.95%) and TikTok (32.55%), followed by YouTube (19.78%), Twitter/X (5.58%) and Facebook (4.14%). This distribution suggests that visual and short video platforms dominate youth online activity, aligning with trends observed in recent Indonesian digital reports (We Are Social, 2024). Consequently, social media serves as a strategic medium for government communication targeting younger audiences.

B. Descriptive Statistics of Research Variables

The descriptive analysis provided insights into participants' perceptions regarding the key variables of the study.

(a) Cultural Nuances in Translation (X1)

The overall mean score for cultural nuances was 3.62 (SD = 0.92), indicating a moderately positive perception. Respondents generally agreed that translations of government content retained important Indonesian cultural expressions and historical references. However, explanatory elaboration for cultural terms received relatively lower agreement (M = 3.24), suggesting that contextual clarification remains limited. This finding implies that while government translations maintain symbolic authenticity, they may lack cultural explication strategies that enhance understanding across diverse audiences.

TABLE 1
CULTURAL NUANCES IN TRANSLATION (X1)

Code	Items	Mean	SD	Category
X1_1	The translation retains uniquely Indonesian cultural terms	3.73	0.91	Positive
X1_2	The translation clarifies the meaning of foreign cultural terms.	3.24	0.94	Neutral
X1_3	The translation's style is appropriate to the cultural context.	3.58	0.90	Positive
X1_4	The translation preserves historical-cultural symbols/references.	3.68	0.87	Positive
X1_5	The translation preserves the emotional nuances of the culture.	3.68	0.87	Positive
Average X1		3.62	0.92	Positive

(b) Translation Strategies (STR)

The mean for translation strategies was 3.47 (SD = 0.84), indicating moderate implementation of cultural translation approaches. While respondents observed domestication and dynamic equivalence strategies, the lower mean for explanatory notes (M = 3.43) implies that translators seldom supplement texts with clarifying annotations. This limited visibility of cultural mediation techniques may contribute to partial comprehension among readers.

TABLE 2
TRANSLATION STRATEGIES (STR)

Code	Items	Mean	SD	Category
STR_1	Translation makes foreign terms sound familiar (domestication)	3.58	0.86	Positive
STR_2	Translation retains foreign elements (foreignization)	3.27	0.83	Neutral
STR_3	Focuses on meaning and message effect (dynamic)	3.53	0.84	Positive
STR_4	Maintains formal government style	3.52	0.81	Positive
STR_5	Provides additional notes/explanations	3.43	0.85	Positive
Average STR		3.47	0.84	Positive

(c) *Machine Translation (MT)*

The construct of machine translation yielded a mean of 3.77 (SD = 0.94), reflecting respondents' awareness of the increasing use of automated translation tools in government communication. They expressed greater trust in human-edited translations (M = 3.84), which they perceived as more reliable in preserving cultural nuances. This finding aligns with global concerns that machine translation, despite its efficiency, often overlooks socio-pragmatic subtleties crucial to public trust.

TABLE 3
MACHINE TRANSLATION (MT)

Code	Items	Mean	SD	Category
MT_1	Much government content is machine translated.	3.79	0.93	Positive
MT_2	Auto-translation doesn't capture cultural nuances well.	3.60	0.95	Positive
MT_3	Human translation is more culturally sound.	3.84	0.91	Positive
MT_4	I've encountered errors in meaning due to machine translation.	3.86	0.95	Positive
MT_5	I trust human translation more than machine translation.	3.77	0.94	Positive
Average MT		3.77	0.94	Positive

(d) *Socio-Political Context (CTX)*

The mean score for socio-political context was 3.52 (SD = 0.84). Respondents indicated that political stability, public trust, and social media discourse significantly shape how they interpret government messages. This suggests that message comprehension is not only influenced by linguistic clarity but also by contextual and affective factors, consistent with communication models emphasizing contextual framing (Criado & Villodre, 2022).

TABLE 4
SOCIO-POLITICAL CONTEXT (CTX)

Code	Items	Mean	SD	Category
CTX_1	Messages are better received when the context is stable.	3.43	0.85	Positive
CTX_2	Social/political issues influence interpretation.	3.54	0.82	Positive
CTX_3	Trust in government influences understanding.	3.62	0.86	Positive
CTX_4	Public reaction on social media influences interpretation.	3.71	0.87	Positive
CTX_5	Identity/cultural issues influence message reception.	3.31	0.82	Positive
Average CTX		3.52	0.84	Positive

(e) *Message Comprehension (X2)*

Message comprehension averaged 3.58 (SD = 0.85). Respondents reported understanding the general meaning of translated government messages but perceived limitations in grasping implicit meanings. This suggests that translations succeed at conveying literal content but may fall short in transmitting nuanced, culture-bound messages that require deeper interpretation.

TABLE 5
MESSAGE COMPREHENSION (X2)

Code	Items	Mean	SD	Category
X2_1	Government messages are easy to understand	3.46	0.83	Positive
X2_2	Translations help capture the implied meaning	3.63	0.85	Positive
X2_3	Messages feel relevant to my life	3.49	0.84	Positive
X2_4	I am able to explain the essence of the message	3.69	0.86	Positive
X2_5	I trust the message if the translation is clear and contextual	3.62	0.87	Positive
Average X2		3.58	0.85	Positive

(f) *Digital Literacy (Y)*

Digital literacy scores across the UNESCO dimensions showed variability. The highest dimension was Ethics & Safety (M = 4.10), followed by Access & Navigation (M = 3.66) and Information Evaluation (M = 3.58), while Creation & Participation (M = 3.06) scored lowest. These findings indicate that Indonesian youth demonstrate strong ethical awareness but limited participatory and creative engagement. This imbalance highlights the need for interventions that move beyond access and safety toward critical and participatory literacy.

TABLE 6
DIGITAL LITERACY (Y)

Dimensions	Item Code	Mean	SD	Category
Access & Navigation	Y_Akses1, Y_Akses2	3.66	0.91	Positive
Information Evaluation	Y_Eval1, Y_Eval2	3.58	0.83	Positive
Creation & Participation	Y_Kreasi1, Y_Kreasi2	3.06	1.06	Negative
Ethics & Safety	Y_Etika1, Y_Etika2	4.10	0.94	Very Positive
Average Y		3.60	0.94	Positive

Overall, youth demonstrated stronger competencies in digital access and ethical use, while their participation and content creation remained underdeveloped.

C. Reliability Analysis

All constructs exhibited Cronbach's Alpha values above 0.70, confirming internal consistency. Machine Translation ($\alpha = 0.88$) and Digital Literacy ($\alpha = 0.89$) showed particularly high reliability, indicating robust measurement across items.

TABLE 7
RELIABILITY ANALYSIS

Variable	Mean	SD	Cronbach's α
Cultural Nuances (X1)	3.62	0.92	0.84
Translation Strategies (STR)	3.47	0.84	0.82
Machine Translation (MT)	3.77	0.94	0.88
Socio-Political Context (CTX)	3.52	0.84	0.79
Message Comprehension (X2)	3.58	0.85	0.83
Digital Literacy (Y)	3.60	0.94	0.89

V. DISCUSSION

The findings demonstrate that cultural nuances crucially shape how young audiences interpret translated government messages, reinforcing Venuti's (1995) concept of translation as a cultural act rather than purely linguistic. Moderately positive perceptions of cultural nuances ($M = 3.62$) suggest Indonesian government translations preserve symbolic and emotional elements of national identity—such as gotong royong and Pancasila—contributing to authenticity and trust. However, the lower score for explanatory elaboration indicates limited contextual clarification. This aligns with Schäffner's (2012) argument that translation without cultural mediation risks partial understanding and weak engagement. It also resonates with Arafah and Hasyim (2023b) who argue misunderstanding in digital communication often arises from neglect of cultural symbols and contextual connotations. In government communication, similar issues occur when translations fail to explain culturally embedded meanings, reducing engagement.

The mediating analysis reveals translation strategies modestly enhance comprehension. Though domestication and dynamic equivalence were viewed positively, their limited mediating power ($\beta = 0.12$, $p < 0.05$) suggests current public institution translation approaches may prioritize formality over accessibility. This echoes Nida's (1964) idea that communicative effectiveness relies on dynamic equivalence, where the target audience's response should mirror the source message's intended impact. In this study, strategies prioritizing literal accuracy or bureaucratic tone seem less effective in conveying socio-emotional meaning, especially to younger audiences favoring contextual, relatable communication on digital platforms. This aligns with Arafah et al. (2022) in "Semiotic Multimodality Communication in the Age of New Media", which emphasizes meaning-making in digital contexts depends on linguistic precision and multimodal interaction between text, visuals, and audience experience. Similarly, government translations need visual and emotional cues aligning with the audience's digital culture.

Machine translation emerged as a significant moderating variable, negatively influencing the relationship between cultural nuances and message comprehension. This supports concerns by Kenny et al. (2024) that automated systems often miss pragmatic subtleties and cultural connotations. Respondents' higher trust in human-edited translations confirms human translators are essential for preserving nuance, intent, and rhetorical tone in intercultural government communication. Moreover, the finding highlights the ethical implication of over-relying on AI-generated translations in public information dissemination—especially when cultural misinterpretation could reduce message credibility or reinforce misinformation. Arafah et al. (2023) further illustrate this, showing nonverbal and emotive signs like emojis carry essential paralinguistic meaning that technology-based systems often misinterpret. The absence of such affective and contextual cues in automated translations can lead to detached or mechanical communication in government messages.

Conversely, socio-political context positively moderated comprehension, showing that cultural translation interacts with factors like political trust and societal climate. In stable conditions, government messages with cultural references are more likely interpreted favorably. This aligns with Criado and Villodre (2022), who emphasize that public trust and

contextual framing mediate citizen perception of institutional messages. In Indonesia, where cultural and political legitimacy intertwine, translation effectiveness depends on linguistic accuracy and perceived communicative source integrity. Thus, sociopolitical harmony amplifies the interpretive potential of culturally nuanced messages (Mare et al., 2025).

The relationship between message comprehension and digital literacy offers important insights. Findings indicate that while youth display strong competencies in ethical and safe digital practices, their participation and content creation levels remain low. This pattern reflects a “consumption-oriented” digital literacy (Ng, 2012), where users excel at accessing and evaluating information but rarely engage as contributors. The limited participatory literacy observed ($M = 3.06$) suggests comprehension does not automatically translate into civic engagement. As Livingstone et al. (2021) argue, critical digital literacy entails not only interpreting information but also challenging, creating, and contributing to public discourse. Therefore, even when translation successfully conveys cultural nuances, its transformative impact on digital citizenship may remain constrained without participatory encouragement. Arafah and Hasyim (2023) stress that true digital literacy involves reflective interpretation and empathetic communication rather than passive consumption. Likewise, Arafah et al. (2022) point out that participatory engagement online depends on users’ ability to interpret multimodal signs—text, images, and symbols—critically and creatively. From a theoretical perspective, this study integrates translation studies, intercultural communication, and digital literacy frameworks. By showing that translation is mediated by socio-political trust and digital competence, the study bridges traditionally separate disciplines. This finding expands Venuti’s (1995) cultural model of translation into the digital sphere, illustrating how cultural negotiation occurs between texts, languages, institutions, and citizens online. The study also operationalizes a four-dimensional digital literacy framework—access, evaluation, creation, and ethics—within a sociocultural communication context, offering a novel view on how translation quality shapes youth digital engagement. The interplay of multimodal communication (Arafah et al., 2022) and paralinguistic meaning (Arafah et al., 2023d) highlights that digital comprehension is intertwined with cultural cues, visual literacy, and emotional intelligence—key components of intercultural translation effectiveness.

Results highlight the need for government institutions to adopt context-sensitive translation frameworks combining linguistic clarity with cultural depth. This involves training translators and digital communication officers to employ hybrid strategies—balancing domestication for accessibility with foreignization for authenticity. Editorial oversight should ensure machine-translated outputs undergo human post-editing to restore cultural and pragmatic coherence. Beyond translation, government agencies should promote digital literacy programs emphasizing participatory and creative competencies, encouraging youth to become informed consumers and active co-creators of digital civic discourse.

In sum, this study provides evidence that cultural nuances in translation significantly influence message comprehension and digital literacy among youth. The mediation of translation strategies and the contrasting roles of machine translation and socio-political context suggest effective government communication requires understanding language, culture, and trust. These findings contribute to the discourse on culturally responsive communication in the digital public sphere and call for collaboration between translators, policymakers, and educators to strengthen intercultural digital engagement. This conclusion reinforces the multimodal and semiotic perspectives proposed by Arafah et al. (2022) and Arafah et al. (2023), affirming that culturally embedded and visually expressive communication enhances clarity, empathy, and engagement in the digital sphere.

VI. CONCLUSION

This study examined how cultural nuances in translating Indonesian government social media content influence youth digital literacy, with translation strategies as a mediator and machine translation and socio-political context as moderators. Results show cultural nuances enhance message comprehension and digital literacy. Translations preserving cultural expressions improve trust and understanding, while excessive machine translation weakens interpretive depth. Translation strategies have a limited mediating effect, indicating government communication remains formal and less audience-centered. Socio-political stability enhances message interpretation, suggesting contextual trust supports cultural comprehension. Linguistic, cultural, and contextual dimensions shape citizen engagement with digital government messages.

Theoretically, this research bridges translation studies, intercultural communication, and digital literacy, showing effective translation as cultural negotiation within the digital public sphere. Practically, it calls for context-sensitive translation frameworks, human–AI collaboration in translation review, and digital literacy programs promoting active and critical youth participation. Effective digital government communication requires linguistic precision, cultural empathy, and contextual awareness to foster informed, engaged digital citizens.

ACKNOWLEDGEMENTS

This work was supported in part by a BIMA grand from the Ministry of Higher Education, Science, and Technology of Indonesia (Kemendikisaintek) under the Contract Number: 02209/UN4.22/PT.01.03/2025.

REFERENCES

- [1] Abbas, H., Arafah, B., Rahman, F., Pattu, M. A., Junus, F. G., Fachry, M. E., Hanafiah, W., Tenrisanna, R., & Manaf, A. (2024). Ecological Literacy and Local Wisdom of Australian Aboriginal People in *Welcome to My Country* Written by Laklak Burarrwanga and Family. *Theory and Practice in Language Studies*, 14(5), 1493-1498.
- [2] Arafah, B., & Hasyim, M. (2019). Linguistic Functions of Emoji in Social Media Communication. *Opcion*, 25(24), 558-574.
- [3] Arafah, B., & Hasyim, M. (2023a). Digital Literacy on Current Issues in Social Media: Social Media as a Source of Information. *Journal of Theoretical and Applied Information Technology*, 101(10), 3943-3951.
- [4] Arafah, B., & Hasyim, M. (2023b). Digital Literacy: The Right Solution to Overcome the Various Problems of Meaning and Communication on Social Media. *Studies in Media and Communication*, 11(4), 19-30.
- [5] Arafah, B., Hasyim, M., & Abbas, H. (2025a). Netizens as Readers, Producers, and Publishers: Communication Ethics and Challenges in Social Media. *International Journal of Advanced Computer Science and Applications*, 16(4), 197-208.
- [6] Arafah, B., Hasyim, M., Abbas, H., Khaerana, A. St. A., & Takwa, T. (2025b). The Values of Leadership in the Local Wisdom of the Bugis People. *WSEAS Transactions on Environment and Development*, 21, 748-762.
- [7] Arafah, B., Hasyim, M., Aman, R. bin, Agussalim, A., & Latief, M. R. A. (2026). Language as a Gateway to Authentic Tourism: Leveraging Bugis-Makassar Expressions in South Sulawesi's Tourism Industry. *Theory and Practice in Language Studies*, 16(2), 597-604.
- [8] Arafah, B., Hasyim, M., Khaerana, A. S. A., Soraya, A. I., & Ismail, A. (2023a). The Digital Culture Literacy of Generation Z Netizens as Readers, Producers and Publishers of Text on Social Media. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3), 112-123.
- [9] Arafah, B., Kaharuddin, Hasjim, M., Arafah, A. N. B., Takwa, & Karimuddin. (2022). Cultural Relations Among Speakers of South Halmahera Languages. *Theory and Practice in Language Studies*, 13(1), 168-174.
- [10] Arafah, B., Rofikah, U., Kaharuddin, A., Room, F., Takwa, & Purwarno, P. (2023b). Evaluating the Suitability of Printed Materials in the Context of Online Learning. *XLinguae*, 16(2), 16-32.
- [11] Arafah, B., Room, F., Suryadi, R., B., L. O. M. I. H., Juniardi, Y., & Takwa. (2023c). Character Education Values in Pullman's The Golden Compass. *Journal of Language Teaching and Research*, 15(1), 246-254.
- [12] Arafah, B., Sirajuddin, J., Thayyib, M., Room, F., Takwa, & Anayati, W. (2023d). Emotional Management of Defoe's *Robinson Crusoe's* Main Character. *Journal of Language Teaching and Research*, 14(5), 1414-1421.
- [13] Arafah, B., Hasyim, M., Arafah, A. N. B., Taqdir, Nursidah, Ruslin, Arafah, F. R. B. (2025c). Local Wisdom and Leadership Values in Bugis Governance: Insights into Indigenous Political Ethics in South Sulawesi. *Scientific Culture*, 11(3.2), 224-234.
- [14] Arifin, M. B., Arafah, B., & Kuncara, S. D. (2022). Dayak's Sociocultural Situation Through Locality in Lumholtz's *Through Central Borneo* Travel Writing. *Theory and Practice in Language Studies*, 12(12), 2695-2703.
- [15] Arnawa, I. G. N. E. V. & Arafah, B. (2023). Students' Self-Regulated Strategies in Approaching Second Language Writing. *Theory and Practice in Language Studies*, 13(3), 690-696.
- [16] Asri, D., Arafah, B., Sahib, H., & Abbas, H. (2023). Male Domination in Helen Garner's *Monkey Grip*. *Theory and Practice in Language Studies*, 13(7), 1651-1658.
- [17] Baa, S., Wardani, S. B., Iskandar, Weda, S., & Arafah, B. (2023). Lexical Metaphors in Westlife's Selected Song Lyrics. *XLinguae*, 16(1), 132-154.
- [18] Bonsón, E., Torres, L., Royo, S., & Flores, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. *Government Information Quarterly*, 29(2), 123-132.
- [19] Criado, J. I., & Villodre, J. (2022). Revisiting social media institutionalization in government. An empirical analysis of barriers. *Government Information Quarterly*, 39(2), 1-13.
- [20] Da Silva, K. C. (2014). Discourse 2.0: Language and new media. *New Media & Society*, 16(4), 696-698.
- [21] Effendy, C., Kurniawan, S., Arafah, B., & Hidayat, S. (2025). Ghosts in Pontianak Malay Beliefs: Origins, Teachings, and Evolving Functions. *Scientific Culture*, 11(3.1), 1476-1492.
- [22] Halil, N. I., Arafah, B., Saputra, I. G. P. E., Hasyim, R. S., Sarmadan, Takwa, & Karma, R. (2024). Preservation of Tolaki Mekongga Language Through Merdeka Curriculum-Based Local Subject Teaching Modules. *Journal of Language Teaching and Research*, 15(3), 960-971.
- [23] Hall, E. T. (1976). *Beyond Culture*. New York: Anchor Press.
- [24] Hasjim, M., Arafah, B., Kaharuddin, Verlin, S., & Aziz Genisa, R. A. (2020). Principles behind Semantic Relation between Common Abbreviations and their Expansions on Instagram. *International Journal of Criminology and Sociology*, 9, 2270-2276.
- [25] Hasjim, M., Arimi, S., Kushartanti, B., Arafah, B., & Taqdir. (2026). Language Transformation in Social Media and Its Impact on the Linguistic Identity of Indonesian Youth. *Journal of Language Teaching and Research*, 17(1), 296-305.
- [26] Hasyim, M., & Arafah, B. (2023a). Semiotic Multimodality Communication in The Age of New Media. *Studies in Media and Communication*, 11(1), 96-103.
- [27] Hasyim M., & Arafah, B. (2023b). Social Media Text Meaning: Cultural Information Consumption. *Wseas Transactions on Information Science and Applications*, 20(2023), 220-227.
- [28] Hatim, B., & Mason, I. (2014). *Discourse and the Translator*. Routledge.
- [29] House, J. (2015). *Translation as Communication across Languages and Cultures*. Routledge.
- [30] Idhan, M., Jabir, Muh., Ruslin, R., Asmawi, M. N., Arafah, B., Ubadah, Ubay, & Fatima. (2026). Teachers' Experiences With Mobile Apps for Arabic Language Learning: Insights From Islamic Schools in Palu, Indonesia. *Journal of Language Teaching and Research*, 17(1), 338-350.
- [31] Iksora, Arafah, B., Syafruddin, S., Muchtar, J., & Lestari, P. A. (2022). Typos' Effects on Web-Based Programming Code Output: A Computational Linguistics Study. *Theory and Practice in Language Studies*, 12(11), 2460-2469.
- [32] Jamiluddin, Suparti, Arafah, B., & Marhum, M. (2026). Language, Power, and Symbolic Violence in Classroom Communication: Insights From Indonesian Culture. *Journal of Language Teaching and Research*, 17(2), 735-746.

- [33] Kaharuddin, Ahmad, D., Mardiana., Latif, I., Arafah, B., & Suryadi, R. (2024). Defining the Role of Artificial Intelligence in Improving English Writing Skills among Indonesian Students. *Journal of Language Teaching and Research*, 15(2), 568-678.
- [34] Kaharuddin, Arafah, B., Ahmad, D., Nursidah, Taqdir, Mardiana, & Rasyid, M. N. A. (2025). Effect of Interactive and Collaborative Strategies on Indonesian EFL Learners' Oral English Proficiency Improvement. *Journal of Language Teaching and Research*, 16(5), 1628-1639.
- [35] Kaharuddin, Arafah, B., Nurpahmi, S., Sukmawaty, Rahman, I. F., & Juniardi, Y. (2023). Exploring How Reading Aloud and Vocabulary Enrichment Shape English Speaking Skills Among Indonesian Learners of English. *World Journal of English Language*, 13(8), 436-445.
- [36] Kenny, Dorothy, & Winters, M. (2024) Customization, Personalization and Style in Literary Machine Translation. In Marion Winters, Sharon Deane-Cox, & Ursula Böser (Eds.), *Technological Change in Translation and Interpreting* (pp. 59-79). Bloomsbury Advances in Translation, London: Bloomsbury.
- [37] Kuswanti, W. H., Arafah, B., Budiman, A. N. A., Ali, T., Fatsah, H., & Room, F. (2023). Students' Perception of Explicit and Implicit Methods in Learning Tenses in SMP DDI Mangkoso. *Theory and Practice in Language Studies*, 13(6), 1473-1482.
- [38] Manugeran, M., Arafah, B., Purwarno, P., Siwi, P., Ekalestari, S., & Wulan, S. (2023). An Ecoliterature Approach to Environmental Conservation: Take Four Selected Literary Works as Examples. *Theory and Practice in Language Studies*, 13(5), 1318-1327.
- [39] Mare, B. S., Arafah, B., Abbas, H., Rano, W. A., Almahdali, U. H., Wicaksono, C. K., Sari, R. A., & Malik, N. F. (2025). The Influence of Social Norms and Etiquette on the Behavior and Decision-Making of Characters in Jane Austen's *Pride and Prejudice*. *Theory and Practice in Language Studies*, 15(12), 4006-4016.
- [40] Marwaha, J. S., & Kvedar, J. C. (2021). Cultural adaptation: a framework for addressing an often-overlooked dimension of digital health accessibility. *NPJ Digit Med.*, 4(1), 143-151.
- [41] Misnah, M., Arafah, B., Bahri, B., Usman, S., Ratu, B., Nurwahyuni, N., & Ruslin, R. (2024). Lore Lindu Culture-Based Education Learning Development for Elementary School Students. *International Journal of Learning, Teaching and Educational Research*, 23(6), 640-661.
- [42] Mofu, H., Arafah, B., & Malawat, I. (2024). Ethnolinguistic Study of Marine Fishes' Characters of the Biak Tribe, Papua. *Theory and Practice in Language Studies*, 14(8), 2532-2542.
- [43] Mokoginta, K., & Arafah, B. (2022). Negotiation in Indonesian Culture: A Cultural Linguistic Analysis of Bahasa Indonesia Textbooks. *Theory and Practice in Language Studies*, 12(4), 691-701.
- [44] Mutmainnah, Arafah, B., & Pattu, A. (2022). Racial Discrimination Experienced by Black People as Reflected in Langston Hughes's Poems. *Journal of Language Teaching and Research*, 13(2), 350-356.
- [45] Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065-1078.
- [46] Nida, E. A. (1964). *Toward a Science of Translating*. Leiden: Brill.
- [47] Nittas, V., Daniore, P., Chavez, S. J., & Wray, T. B. (2024). Challenges in implementing cultural adaptations of digital health interventions. *Communications Medicine*, 4(1), 1-7.
- [48] Radjaban, R. Y., Humanika, E. S., & Arafah, B. (2025). A Holistic Study of Merap Language Mapping in Relation to the Term Punan Merap. *Journal of Language Teaching and Research*, 16(3), 957-965.
- [49] Rano, W. A., Arafah, B., Abbas, H., Almahdali, U. H., Wicaksono, C. K., Sari, R. A., Malik, N. F., & Mare, B. S. (2026). Social Class and Stratification in the Edwardian Era as Portrayed in Forster's *A Room With A View*. *Theory and Practice in Language Studies*, 16(2), 639-648.
- [50] Schäffner, C. (2012). *Political Discourse, Media and Translation*. Cambridge Scholars Publishing.
- [51] Shifman, L. (2013). *Memes in Digital Culture*. The MIT Press. <https://doi.org/10.7551/mitpress/9429.001.0001>
- [52] Suhadi, J., Arafah, B., Makatita, F. P., Abbas, H., & Arafah, A. N. B. (2022). Science and Society: The Impact of Science Abuse on Social Life in Well's *The Invisible Man*. *Theory and Practice in Language Studies*, 12(6), 1214-1219.
- [53] Sunardi, Akil, M., Arafah, B., & Salija, K. (2018). Looking at the Shared Conception of Teaching Literature in an Indonesian ELT Setting. *Journal of Language Teaching and Research*, 9(2), 316-327.
- [54] Suparti, Kholiq, A., Arafah, B., Kaharuddin, Suhartono, & Milawati. (2025). The Effect of Brainstorming and Free Writing on Writing Skill Development: Teachers' Feedback as a Mediator. *Theory and Practice in Language Studies*, 15(11), 3785-3797.
- [55] Sunyoto, F. G., Arafah, B., Yudith, M., Mokodompit, G. P., & Asnawi, A. E. F. (2022). The Native American Father's Parenting Style in John Ernst Steinbeck's *The Pearl*. *Theory and Practice in Language Studies*, 12(12), 2551-2558.
- [56] Takwa, Arafah, B., Hasyim, M., & Akhmar, A. M. (2024a). Cultural Imagery of Tolaki Mekongga Community of Kolaka in Mowindahako Ritual. *Theory and Practice in Language Studies*, 14(3), 763-770.
- [57] Takwa, Arafah, B., Hasyim, M., Akhmar, A. M., & Taqdir. (2025). Cultural Value in Mowindahako Ritual Speeches of Tolaki Mekongga Ethnic Group in Kolaka, Indonesia. *Theory and Practice in Language Studies*, 15(6), 1857-1866.
- [58] Takwa, Arafah, B., Hasyim, M., Akhmar, A. M., & Taqdir. (2026). Linguistic Units and Cultural Meanings in the Ritual Speech of Mowindahako Among the Tolaki Mekongga Ethnic Group. *Journal of Language Teaching and Research*, 17(2), 665-676.
- [59] Takwa, Arafah, B., Kaharuddin, Putra, E., Masrur, & Arafah, A. N. B. (2022a). The Shift of Lexicon in Traditional Technology System in Tolaki Community at Konawe District of Southeast Sulawesi. *Theory and Practice in Language Studies*, 12(5), 980-989.
- [60] Takwa, Arafah, B., Sopiandy, D., Taqfiah, S. J., Arafah, A. N. B. (2022b). Humanistic Values in Metaphoric Expressions of Traditional Marriage in Tolaki Mekongga Kolaka. *Theory and Practice in Language Studies*, 12(8), 1602-1608.
- [61] Takwa, T., Arafah, B., Hasyim, Muhammad, Akhmar, A. M., Taqfiah, S. J., & Sarmadan, S. (2024b). Symbolic meanings in the Mowindahako ritual of the Tolaki Mekongga Kolaka ethnic. In J. Jaiswal (Ed.), *International Conference on Medical Imaging, Electronic Imaging, Information Technologies, and Sensors (MIEITS 2024)* (p. 38). SPIE.
- [62] Takwa, T., Arafah, B., Taqdir, T., Sarmadan, S., Tundreng, S., Pageyasa, W., Hariyadi, S., Rasid, R., Inpresani, W. O., & Mofu, H. (2026). Local Knowledge for Coastal Resilience: Adaptive Strategies of Bajo Fishers in Pomalaa, Indonesia. *Scientific Culture*, 12(1.1), 1336-1347.

- [63] Taqdir, Arafah, B., Nursidah, & Afisha A, W. O. N. (2025). Gender Representation in Foreign Language Textbooks: Trends, Gaps, and Implications for Multilingual Teaching. *XLinguae*, 18(4), 225–242.
- [64] Taqdir, Arafah, B., & Nursidah. (2026). Perceptions of Cultural Representation in Japanese Language Textbooks: Insights From Students and Teachers in Indonesian Higher Education. *Journal of Language Teaching and Research*, 17(2), 622–633.
- [65] Taqdir. (2025). Cultural Stereotypes in Foreign Language Textbooks: A Systematic Review of Visual Representation. *Theory and Practice in Language Studies*, 15(6), 1949–1958.
- [66] Usman, S., Arafah, B., Marhum, M., Munir, S., Budi, & Tadeko, N. (2024). The Effect of Technological Pedagogical Content Knowledge (TPACK) on Rural Area Students' English Writing. *Journal of Language Teaching and Research*, 15(3), 884–892.
- [67] Venuti, L. (1995). *The Translator's Invisibility: A History of Translation*. Routledge.
- [68] Venuti, L. (1998). *The Scandals of Translation: Towards an Ethics of Difference*. Routledge.
- [69] Yotha, N. (2025). The Development of a Digital Literacy Assessment Tool for Thai Grade 10-12 Students. *Higher Education Studies*, 15(1), 316-324.
- [70] Yudith, M., Arafah, B., Abbas, H., Jaelani, A., Sunyoto, F. G., & Arafah, A. N. B. (2024). Social Issues During the Victorian Era Lead to the Formation of Nice Girl Syndrome in Gaskell's *Ruth*. *Journal of Language Teaching and Research*, 15(3), 822–832.
- [71] Yudith, M., Arafah, B., Sunyoto, F. G., Fitriani, Rostan, R. B., & Nurdin, F. E. (2023). The Representation of Animalism Issue in Sewell's *Black Beauty*. *Theory and Practice in Language Studies*, 13(1), 108-116.
- [72] Yulianti, S., Arafah, B., Rofikah, U., Idris, A. M. S., Samsur, N., & Arafah, A. N. B. (2022). Conversational Implicatures on Saturday Night Live Talk Show. *Journal of Language Teaching and Research*, 13(1), 189-197.



Burhanuddin Arafah obtained his PhD in English literature at the University of Newcastle, Australia, in 2003. He earned his Master's degree in American literature at Gadjah Mada University (UGM) Yogyakarta of Indonesia in 1995 and his Bachelor's degree in English literature at Hasanuddin University (UNHAS) in 1988. He is currently a full professor of English literature at the English Literature Study Program, Faculty of Cultural Sciences of UNHAS, Indonesia. His areas of interest are English literature, language education, and cultural studies.



Herawaty Abbas obtained her Ph. D. at the University of Newcastle, Australia, in 2013, and her Master's degrees from Gadjah Mada University (UGM) Yogyakarta of Indonesia in 1997 and Saint Mary's University Canada in 2001. Her doctorate thesis examines the potential dialogue between Australian and Buginese cultures regarding feminism. Her research interests are feminist, children's, and Indigenous literature.



Muhammad Hasyim is a professor in the field of linguistics (semiotic studies). He lectures Doctoral Linguistic Study Program, at the Faculty of Cultural Sciences, Hasanuddin University. He completed his doctoral study in linguistics at Hasanuddin University, where he did his dissertation research in semiotics. He has done much research in language, culture and media studies with a semiotic approach.



Muhammad Ridha Anugrah Latief earned his Master's degree (*Magister Humaniora- M.Hum.*) in English Language Studies (English Translation Study) from Hasanuddin University, Indonesia, in 2020. He has worked as a lecturer at English Literature Department, Hasanuddin University since 2020. He currently serves as vice chairperson of HPI Chapter South Sulawesi for 2023 – 2028 period.



Taqdir completed his Master's degree in Literature with a concentration in Japanese Linguistics in 2013 at Padjadjaran University. He is pursuing his doctoral degree in the Linguistics Department at Hasanuddin University. His research focuses on Japanese language and culture. He also actively publishes scientific articles in both national and international journals.

Gendered Variations in Translanguaging Practices and Attitudes by EFL Teachers in Saudi Arabia

Sita N. Aldossari

College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia

Manar Almanea*

College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia

Abstract—This qualitative study investigates gendered variations in translanguaging practices thoroughly among six Saudi secondary school EFL teachers (three males and three females). Using classroom observations followed by retrospective, semi-structured interviews, the study examines the frequency, functions, and underlying attitudes shaping the intended use of Arabic (L1) in English (L2) instruction. The findings reveal notable gender-related differences in translanguaging practices. Male teachers employed Arabic more frequently during instruction, primarily for direct explanation and comprehension checks, whereas female teachers used translanguaging more selectively for scaffolding and rapport-building, adjusting their use according to students' proficiency levels. Despite these differences, both groups maintained English-only instruction during productive skills assessments. The study highlights teacher gender as a significant variable in translanguaging research and advocates for gender-sensitive teacher education that legitimizes strategic L1 use as a culturally responsive pedagogical resource.

Index Terms—translanguaging, EFL instruction, teacher gender, bilingual education, classroom discourse

I. INTRODUCTION

The teaching of English has experienced a significant paradigm shift from monolingual "English-only" ideologies toward embracing multilingual approaches that recognize the pedagogical value of students' full linguistic repertoires (Canagarajah, 2011; Cenoz & Gorter, 2022). Translanguaging has been recently recognized as an inclusive pedagogical practice that promotes language equality by validating learners' linguistic identities and challenging monolingual ideologies that marginalize minoritized languages in EFL classrooms (García & Wei, 2014; García et al., 2017). It refers to the intended integration of multiple linguistic resources that has emerged as both a purposeful pedagogical method and an instinctive communicative practice in diverse educational contexts (García & Otheguy, 2020). In Saudi Arabia, this shift aligns with the goals of Vision 2030, within which the Human Capability Development Program identifies English proficiency as a national priority (Almanea et al., 2024; Vision 2030, 2021). Recent research demonstrates that pedagogical translanguaging yields beneficial results in Saudi classrooms, notably in reducing student anxiety, enhancing comprehension, and strengthening newly bilingual learners (Alsagoafi, 2025). However, ongoing gaps persist between policy goals and classroom practices (Alqahtani, 2022), particularly due to earlier EFL teachers' education, which insisted on an English-only medium of instruction. Enduring concerns persist about potential restrictions on Arabic use for English proficiency development (Alzahrani & Althaqafi, 2023; Alasmari et al., 2022).

Despite widespread adoption of translanguaging in Saudi EFL classes, significant gaps remain in comprehending the influence of teacher gender on these linguistic practices. Although current studies have outlined general code-switching patterns and instructional roles among Saudi educators (Alshammari, 2011; Alsagoafi, 2025), essential questions remain: Do male and female teachers engage in translanguaging at different frequencies? Do they utilize Arabic for different educational objectives? Do they possess differing perspectives on L1 utilization in English instruction? These questions hold particular significance in Saudi educational settings, where gender-segregated classrooms may produce distinctly gendered translanguaging patterns. The Kingdom's distinctive sociolinguistic context—marked by changing gender norms, historically separated education, and differing societal expectations for male and female teachers—provides a substantial foundation for examining the influence of gender on translanguaging practices (Al-Rasheed, 2021).

This study addresses both theoretical and practical gaps by examining gendered patterns of translanguaging among Saudi EFL teachers, focusing on observed classroom practices and underlying attitudes toward L1 utilization in English instruction. Despite significant research on translanguaging in many educational contexts globally, the connection between these practices and gender is insufficiently explored, specifically in the Saudi context. Recent research has documented gender differences in perceptions of translanguaging among EFL teachers in a different context, with female

* Corresponding Author. Email: manar_ut@yahoo.com

teachers demonstrating more favorable attitudes toward multilingual pedagogical approaches (Ulum, 2024). Indeed, research on classroom interactions demonstrates that male and female teachers often adopt different communication styles, with female teachers engaging in more nurturing and collaborative approaches and male teachers adopting more authoritative communication patterns (Einarsson & Granström, 2002; Sadker et al., 2009). In a study by Alsagoafi (2025), strategic translanguaging, which improved student performance and self-esteem in secondary schools, required a shared linguistic grounding between teachers and students. This indicates that effective translanguaging practice depends upon contextual elements that remain inadequately investigated. As Saudi Arabia navigates the balance between tradition and modernization through Vision 2030 educational reforms (Al-Rasheed, 2021), comprehending these gendered linguistic practices is crucial for formulating culturally responsive pedagogies that recognize both multilingual realities and evolving national identity. Accordingly, this study addresses three interrelated questions:

1. What patterns of translanguaging between Arabic and English do male and female Saudi EFL teachers display in their classroom instruction?
2. How do the functions of translanguaging practices differ between male and female Saudi EFL teachers?
3. What attitudes do male and female Saudi EFL teachers hold toward the use of Arabic in English language classrooms?

II. LITERATURE REVIEW

Recent research in multilingual discussions has already shifted from focusing on 'natural' code-switching theories, which view bilingualism as the alternation between different language codes (Myers-Scotton, 1993), to 'intentional' translanguaging frameworks that acknowledge the integrated linguistic collections of multilingual speakers. Translanguaging refers to "the fluid, dynamic practices that transcend the boundaries between named languages, language varieties, and between language and other semiotic systems" (Li, 2018, p. 9), challenging traditional understandings of bilingual language use as the switching between separate codes. García and Otheguy (2020) further argue that multilingual individuals draw on their full linguistic repertoire in response to communicative requirements, rather than perceiving language use as alternating between separate structures. Cenoz and Gorter (2022) emphasize that this redefinition has significant educational consequences: translanguaging views students' full linguistic competencies as significant resources for comprehension and learning, rather than treating L1 use as a deviation from target-language norms (Vogel & García, 2021).

Recent research demonstrates that strategic L1 use in language instruction yields modest yet consistent benefits across multiple dimensions, including content comprehension, metalinguistic awareness, and affective outcomes (Turnbull & Dailey-O'Cain, 2009). Macaro's (2024) comprehensive review confirms that cautious L1 use facilitates understanding of cognitively demanding content, particularly for lower-proficiency learners, without hindering target language vocabulary acquisition. Similarly, Cenoz and Gorter (2022) found that translanguaging-supportive educational environments enhanced students' metalinguistic awareness and fostered more positive attitudes toward minority languages compared to language-separated programs.

English education in Saudi Arabia is embedded in sophisticated sociolinguistic frameworks, where English primarily serves as an educational and professional tool, while Arabic dominates social and cultural domains (Elyas & Badawood, 2016). Alasmari et al. (2022) examined university-level EFL courses and found that instructors used Arabic with moderate frequency ($M = 3.50$ on a 5-point scale) across various classroom practices, primarily for explaining concepts, giving directions, and managing the classroom. However, most instructors expressed disapproval, describing the use of Arabic as necessary scaffolding that should ideally be reduced, reflecting persisting monolingual ideals despite practical situations. In a recent study, Alsagoafi (2025) found that secondary school teachers strategically employed translanguaging to scaffold, particularly during brainstorming and when addressing communication breakdowns, with reported improvements in student performance and self-esteem among lower-proficiency learners. EFL teachers utilized translanguaging across different pedagogical stages while teaching different skills. In writing classes, teachers emphasized translanguaging during brainstorming to ensure clarity of the task. In reading activities, teachers applied translanguaging through scaffolding to support comprehension, particularly with difficult vocabulary. Teachers reported that translanguaging made content more accessible and improved student engagement, especially for lower-proficiency learners. Critically, successful implementation occurred because teachers and students shared Arabic and English, enabling fluid language switching. Apparently, these previous studies generally disregard gender as a crucial variable, leaving essential questions unresolved about potential differences between male and female teachers' practices, particularly in Saudi Arabia's gender-separated educational system.

Teachers occupy distinctive positions as translanguaging agents, simultaneously constrained by institutional policies and empowered as decision-makers within classrooms. Their practices reflect pedagogical judgments, professional identities, and beliefs about effective teaching (Canagarajah, 2011). Alturki and Alharbi (2024) found that Saudi teachers construct professional identities differently by gender and career stage, thereby affecting their linguistic choices in classrooms. Saudi female teachers predominantly framed their roles in terms of caring metaphors (mothers and mentors), whereas male teachers emphasized knowledge transmission and achievement. These varying identity constructions reflect societal expectations and institutional positioning: teachers who view themselves primarily as language models may minimize L1 use, whereas those emphasizing mentoring roles may use translanguaging more frequently to build rapport.

In Saudi Arabia, gender intersects with religion, cultural traditions, and rapid social change, producing complex sociolinguistic patterns (Al-Wer et al., 2022). Recent research demonstrates that Arab women negotiate multiple, often conflicting expectations through strategic language use, with educated women frequently leading linguistic innovation while maintaining elements signaling cultural authenticity (Bassiouny, 2020). Tensions between traditional gender norms emphasizing family roles and contemporary aims for professional achievement significantly shape female teachers' professional practices (Al-Rasheed, 2010). In gender-segregated educational contexts, female teachers are expected to provide not only academic instruction but also moral guidance and emotional support, which may influence their linguistic choices. These expectations position women at the intersection of tradition and modernity, requiring strategic negotiation through language use (Al-Wer et al., 2022).

Emerging research suggests that gender may significantly influence how teachers implement translanguaging, yet systematic research on this topic remains limited. Studies across various bilingual contexts indicate that female teachers may use L1 more frequently for affective purposes, such as building rapport and managing anxiety, whereas male teachers tend to restrict L1 use primarily to content clarification (Litosseliti, 2006). These patterns may reflect broader gendered communication styles, with women engaging in more relational work and men focusing on transactional communication (Holmes & Schnurr, 2006). Critical approaches to translanguaging examine how power relations and institutional ideologies shape who engages in translanguaging, under what conditions, and with what forms of support, revealing important equity concerns in multilingual education (Paulsrud et al., 2017; Lin & He, 2022; Conteh & Brock, 2021).

Within Saudi contexts, Alharbi and Alqefari (2022) investigated university teachers' perceptions of translanguaging through surveys and semi-structured interviews with five professors, finding that teachers favored translanguaging because they shared linguistic and cultural backgrounds with students. However, their study did not systematically analyze gender differences despite including both male and female participants. Alsagoafi's (2025) recent study provides the most direct evidence concerning gender and translanguaging in Saudi EFL contexts. While the study collected data from equal numbers of male and female secondary teachers (five each) and conducted interviews with both groups (three women and three men), the published findings do not systematically compare gender differences. The study reports general patterns that teachers used translanguaging for scaffolding, to explain difficult words, to encourage student interaction, and to support lower-proficiency learners.

Existing Saudi research has not systematically analyzed potential gender differences in translanguaging practices. Moreover, the reviewed studies rely primarily on self-report data from questionnaires and interviews rather than systematic classroom observation. This study addresses this gap by (1) systematically analyzing gender as a primary variable rather than treating teachers as a homogeneous group, (2) integrating classroom observation with teacher interviews to examine alignment between practices and attitudes, and (3) investigating factors shaping gendered translanguaging patterns within Saudi Arabia's specific sociocultural context.

III. METHODOLOGY

A. Research Design

This study employed a qualitative research design to investigate gendered variations in translanguaging practices among Saudi EFL secondary school teachers. Data were collected through first classroom observations followed by in-depth semi-structured interviews. Then analyzed using Braun and Clarke's (2012) six-phase thematic analysis framework.

B. Participants

Six English language teachers (three females and three males) were purposefully selected from public secondary schools in Saudi Arabia. All participants held bachelor's degrees in English from Saudi universities. All six participants were experienced EFL teachers at the secondary level, with 4 to 11 years of experience (Ave.=8). Their ages ranged between 26-34 years old (Ave.=30.2). Their classes include around 33-38 students (Ave.=35). The textbook series is selected by the Saudi Ministry of Education and is taught in all Saudi Schools. It is called *Mega Goal* by McGraw Hill. Female participants were assigned codes FT1, FT2, and FT3, while male participants were coded MT1, MT2, and MT3 to maintain confidentiality.

C. Data Collection

Data were collected through two primary sources: classroom observations and semi-structured interviews. This dual-method approach enabled triangulation, enhancing validity by comparing teachers' reported practices with observed classroom behaviors (Cohen et al., 2018). Classroom observations were conducted by the researchers and a male colleague. They documented naturally occurring translanguaging behaviors in authentic instructional settings. Each observation lasted 45 minutes (total=270 minutes) and employed a structured observation sheet (Appendix A) to systematically record (1) contextual information (teacher gender, grade level, lesson topic), (2) translanguaging functions using predetermined codes: INS (Instructional: explaining grammar, vocabulary, concepts), MGT (Management: organizing activities, giving directions, discipline), AFF (Affective: encouraging, praising, building rapport, humor), CHK (Checking comprehension), COR (Corrective feedback), PER (Personal/Off-topic), and (3) an observation log capturing what was said in Arabic, translanguaging triggers, and student responses. In addition to live observation, lessons were audio-recorded using the researchers' smart phones to allow for further analysis. The percentage of Arabic language

use in these classes was calculated by dividing the number of utterances in Arabic by the total number of utterances in the lesson.

Semi-structured interviews followed classroom observations and lasted 20-30 minutes (total=153 minutes). They were conducted face-to-face in English and audio-recorded to allow later analysis. They consisted of three sections: (1) teaching background and English proficiency, (2) translanguaging practices, including purposes for Arabic use (explaining grammar/vocabulary, managing behavior, giving instructions, building rapport, encouraging students), and (3) beliefs, attitudes, and personal teaching philosophies regarding L1 use in English instruction. The flexible interview format allowed participants to elaborate on issues while ensuring key topics were systematically explored (Cohen et al., 2018). Interviews were transcribed for analysis purposes.

D. Ethical Considerations

Institutional ethical approval was obtained before data collection. Additionally, participants provided informed consent prior to participation, having been fully informed of the study's purpose, research procedures, their rights, and confidentiality measures. The consent form assured protection of participants' privacy and personal data, clarified that participation was voluntary, and explained that participants could withdraw at any time without consequences. Anonymity was maintained using coded identifiers (FT1-FT3 for females; MT1-MT3 for males). All data were stored securely and used solely for research purposes.

E. Data Analysis

Qualitative data were analyzed using Braun and Clarke's (2012) six-phase thematic analysis: (1) familiarizing with data through repeated reading/listening and documenting initial observations, (2) generating initial codes to identify meaningful data segments, (3) searching for patterns and clustering codes into broader categories, (4) reviewing and refining themes, (5) defining and naming themes, and (6) producing the final analytical narrative with illustrative examples. To enhance trustworthiness, multiple strategies were employed, including member checking to verify interpretations and triangulation between interview and observation data to strengthen credibility and confirmability (Braun & Clarke, 2012; Creswell & Creswell, 2022). Both researchers independently coded the data, and their inter-rater reliability was calculated using Cohen's Kappa ($K = 0.78$).

IV. RESULTS

A. Classroom Observation Results

Classroom observations revealed varying translanguaging patterns across the six participants. FT1 used Arabic minimally (approximately 10% of the lesson), primarily for giving specific instructions. The third-grade students' advanced English proficiency enabled them to understand the teacher's questions and instructions without requesting translation. The lesson focused on listening comprehension and oral discussion of the topic. Arabic served functional purposes such as clarifying instructions (codes INS, MGT, and COR; see Appendix B), with examples including "هذا الجدول مهم لفهم حل النشاط" (*This table is important to understand the exercise*) and "سجلوا هذه الملاحظات علشان ما تتكرر نفس الأخطاء." (*Write down these notes so you do not make the same mistakes again*).

During a reading class, FT2 used Arabic for approximately 40% of the lesson with second-grade students. Arabic was primarily used to support comprehension of new vocabulary. It was also used to explain structural distinctions between transitional sentences (TS), supporting sentences (SS), and concluding sentences (CS). Examples include "نفس المدونة، تعنى" *blog means ...* and "تعى السياحة ... *Tourism means ...*" The teacher consistently used the Arabic word "عادة" instead of "usually" throughout the class.

FT3 used Arabic for approximately 60% of the lesson with first-grade students during a speaking lesson titled "What do I need?" Despite this high level of translanguaging, all other aspects of instruction, including general classroom interactions, were conducted in English. Arabic was primarily used to explain new vocabulary and manage classroom activities (codes INS, GMT, and COR; see Appendix B), such as "إيش المشكلة عندك" (*what is your problem?*) and "يعنى" (*the household costs*). The average translanguaging frequency percentage by all female instructors is 36.6%.

MT1 used Arabic for approximately 45% of the lesson with first-grade students during a grammar lesson. Translanguaging most frequently occurred after each question. The teacher employed Arabic selectively to clarify students' responses and reinforce understanding, while the remainder of the lesson was conducted in English. It appeared that the students needed additional practice answering questions without Arabic support. Examples of Arabic usage include "استمع الي النص و ناقش الأسئلة التالية" (*listen to the audio and discuss the following questions*) and "نستخدم الفعل المناسب" (*we use the appropriate verb*) (codes INS, COR, MGT; see Appendix B). Translanguaging also helped relieve time pressure and supported positive teacher-student relationships, particularly when explaining complex grammar points.

MT2 used Arabic for approximately 70% of the lesson with second-grade students during a reading lesson about geographical directions. Students were unable to answer most questions without Arabic translation and struggled to understand the teacher's English instructions. Arabic was extensively used to introduce new vocabulary, clarify grammar points, and manage classroom activities. Examples include "احفظ الصورة و أشر عليها" (*memorize the sign and point to it*) and "ما هي الجهات الأساسية في البوصلة بالعربي" (*what are the primary directions in the Compass in Arabic?*) (codes INS, GMT; see

Appendix B). MT2's strategic use of Arabic humor to create rapport "شكلكم نايمين اليوم مو حاضرين الدرس" (*Are you sleeping, not attending?*) made them smile and relax.

MT3 used Arabic for approximately 78% of the lesson with third-grade students during a grammar lesson explaining adjectives. During the observation, it seemed that students felt comfortable only when the teacher used Arabic. The teacher relied heavily on Arabic to clarify vocabulary, explain complex grammatical rules, and manage classroom activities. Saudi colloquial Arabic was frequently used, such as "ايش تتذكرون من كلمات تدل على ظرف مكاني" (*what words do you remember that refer to place adverbials?*) and "وين مكان الماوس و الكيبورد في الصورة" (*where is the place of the mouse and the keyboard in the picture?*) (codes INS, MGT; see Appendix B). MT3 used Arabic for clarity, immediately rephrasing confusing instructions in Arabic. The average translanguaging frequency percentage by male instructors is 64.3%. Overall, male EFL teachers used translanguaging remarkably more than female teachers.

Table 1 below summarizes the results of the class observations regarding translanguaging frequency. It varied according to teacher gender, grade level, and lesson topic. Both male and female teachers employ higher percentages of translanguaging, particularly with low-proficiency students. It is observed that translanguaging frequency generally rises as students' English proficiency decreases. Grade level interacts with this trend: female teachers reduce translanguaging with higher grades, while male teachers' usage is more dependent on proficiency than on grade level. Lesson content also influences translanguaging frequency, with grammar and reading lessons showing higher percentages than speaking and listening lessons, especially with advanced learners.

TABLE 1
PERCENTAGES OF TRANSLANGUAGING IN THE ATTENDED LESSONS

Gender	Secondary School Grade level	Topic of Lesson	Translanguaging percentage %	Average translanguaging percentage
FT1	3 rd grade	Listening and Discussion	10%	35.6%
FT2	2 nd grade	Reading	40%	
FT3	1 st grade	Discussion	60%	
MT1	1 st grade	Grammar Review	45%	64.3%
MT2	2 nd grade	Reading: Geographical Directions	70%	
MT3	3 rd grade	Grammar: Adjectives	78%	

B. Interview Results

Interviews were coded into themes and subthemes. Table 2 in Appendix C presents the major themes and their interpretations, along with the subthemes for each theme and their interpretations.

Emerging Themes from Interviews

The results revealed five emerging themes. The following sections discuss these themes, supported by illustrative quotations.

(a) Theme 1: Areas of Translanguaging Practices Within Classrooms

Teachers used translanguaging across several instructional and relational domains. Participants reported frequent use of Arabic to clarify grammar and explain new vocabulary. FT1 stated, "When explaining tenses like present perfect, I say 'هو يشبه الماضي لكن مرتبط بالحاضر' "*it resembles the past but is related to the present*" to help them relate it to Arabic structure. Translanguaging shortens explanation time." FT2 noted, "For passive voice, I use Arabic to contrast sentence order. For example, 'في العربية الفاعل يأتي أولاً، لكن هنا المفعول هو الأساس' "*In Arabic, the subject comes first, but here the object is more important.*" It makes it clearer for visual learners." Regarding vocabulary, FT1 said, "Selectively, I translate abstract words like 'justice' *العدالة* when there is no context clue." FT2 noted, "I translate idioms or cultural expressions like 'break a leg' *(بالتوفيق)* to help students understand cultural meaning." However, FT3 pointed out, "I prefer using pictures, synonyms, or examples. Only in rare cases do I provide the Arabic meaning, such as 'honest' *صادق*."

MT1 reported using translanguaging to explain grammar, "Because I believe it will be easier for me to explain the grammatical rule and structures in Arabic than in English, as my students are not very competent in the English language." MT2 noted, "Sometimes Arabic helps clarify complex tenses like present perfect or conditionals. It saves time and ensures comprehension." MT3 noted, "I teach grammar directly. In a recent lesson about self-introductions, students needed to understand the verb 'to be' and how to use pronouns like 'I' and 'my' before they could practice".

Regarding vocabulary, MT1 said, "Every new vocabulary we have in each lesson, I translate it to Arabic." MT2 commented that he uses translanguaging "only when the word has no clear context or visual support. For abstract words like 'envy,' I translate quickly to Arabic, but for concrete nouns, I prefer pictures or examples." MT3 stated, "I focus on high-frequency verbs and vocabulary relevant to their field of study, which is tourism in this case".

Translanguaging also played a key role in managing classroom behavior, giving instructions, and handling difficult situations. FT1 affirmed, "When students are noisy, Arabic works better: 'بنات، خلونا نركز شوي' (*Girls, let's concentrate a bit*). It sounds natural and serious in Arabic." FT3 commented, "I sometimes start in English but end up switching if it is urgent: 'Stop talking... خلاص وقفو كلام'." For giving instructions, FT1 stated, "I usually say it in English first, then repeat key points in Arabic if students seem lost, e.g., 'Work in pairs...' *يعني اشتغلوا كل ثنتين سوا*." FT2 noted, "I use Arabic when instructions are complex: 'اقرأوا القطعة أول بعدين جاوبوا الأسئلة' (*Read the passage first, then answer the questions*). However, FT3 disagreed, saying, "I demonstrate instead of translating. I show them how to do the task rather than saying it in Arabic." MT1 commented that he employed translanguaging "to make it easy for the students to understand and react

immediately." MT2 reported, "When I need quick control, Arabic works better; saying 'اجلسو بسرعة يا شباب' gets attention immediately, while English commands sometimes take longer." He added, "I give instructions (Guys sit down quickly) mainly in English but repeat key points in Arabic if students look confused, especially before exams or group tasks." MT3 added, "I make sure my instructions are clear. If a student seems confused and says, 'The question is not clear,' I immediately rephrase it, often in Arabic".

Participants also used translanguaging to offer positive feedback and encourage and praise students. FT1 said, "I often mix languages: 'Excellent, mashallah!'" FT2 affirmed using Arabic expressions like "يعطيك العافية ما شاء الله تبارك الله يا سارة" and added, "I mostly use the English phrase 'Good work, proud of you!' but sometimes add Arabic to make it personal: 'ممتاز جدا اليوم'". Additionally, participants used Arabic to build rapport and encourage students. FT1 stated, "Arabic jokes make the atmosphere friendly. I might say, 'بس لأن جو الحصة حلو' They laugh and feel connected." (We are not having homework today, not because you were good, but because it was a nice lesson!) FT2 noted, "I use light humor in Arabic to lower anxiety. For example, when a student forgets homework, I say 'ولانتني نسيتي؟ ولا' انتي نسيتي؟" (Did you forget the book? Or did the book forget you?). FT3 reported, "Sometimes I mix languages: 'Good job, wallah, you did amazing!' Students enjoy the blend." MT2 said, "Arabic humor builds a friendly atmosphere. For example, when students are sleepy early in the morning, I might joke in Arabic. That makes them smile and relax." He added, "I mix both languages. I might say 'Excellent work, mashallah!' or 'Great effort, ya Mohammed!' Using Arabic expressions like 'mashallah' feels more personal and culturally natural." MT3 commented that employing translanguaging for joking or building rapport "is an area I am working on. I can be detached from my students, but I see how my colleagues use humor to create a positive classroom atmosphere, and I am trying to adopt similar strategies".

Participants also used translanguaging to support students who were introverted and shy. FT1 said, "For shy students, I switch to Arabic briefly to make them comfortable: 'ما في مشكلة، خذي وقتك' (never mind, take your time) FT2 affirmed that Arabic helps calm anxious students: "لا تتوترى، احنا نتعلم مع بعض" (don't worry, we are together). MT2 said, "I often switch to Arabic with shy or weak students to reduce pressure. It helps them express what they do not understand without feeling embarrassed. For example, I might say 'ما شاء الله رافع بس، حاول مرة ثانية' (Great mashallah, but try again).

These responses reveal that participants employ translanguaging as a multifunctional tool. Arabic is used purposefully to enhance comprehension, manage classroom dynamics, and support instructional clarity. Additionally, translanguaging serves a relational function, helping teachers build rapport, motivate students, and create a positive learning environment.

(b) Theme 2: Situations Where Participants Refuse to Use Arabic

Despite recognizing the benefits of translanguaging, participants identified specific situations in which they intentionally avoided Arabic to preserve English immersion. Participants reported refraining from Arabic during speaking tests, writing lessons, reading introductions, discussion tasks, and listening activities, noting that these contexts require sustained exposure to the target language. Some also avoided Arabic with low-proficiency learners to encourage gradual independence. FT1 reported, "During speaking exams or conversation classes, I insist on English only to build real communication." FT2 said she avoids Arabic in writing lessons. I tell them, 'Think in English.' Switching to Arabic breaks their focus." FT3 noted avoiding Arabic "when I introduce reading passages. I want students to guess meanings in context first before translating".

MT1 stressed not using Arabic specifically for "answering the questions, especially speaking and discussion tasks, because it is essential in practicing the language as an authentic practice." MT2 reported not allowing students to speak in Arabic "during speaking or listening activities," adding, "I insist on using only English. For example, in oral presentations or debates, I tell them, 'English only, please.' It is important to create an English environment to build fluency and confidence." In contrast, MT3 commented, "My current approach is to use Arabic to ensure student comprehension. If I used only English, many students would not be able to understand the lesson. In an ideal situation, using only English would be highly beneficial for their language acquisition, as the classroom is their primary environment to practice. However, this is very difficult to implement effectively when the students' proficiency level is low".

(c) Theme 3: Participants' Strategies for Managing Students' Use of Arabic

Participants described a range of strategies for managing students' spontaneous use of Arabic. Many adopted a selective translanguaging approach by responding partially in Arabic, using brief Arabic reassurance, or resorting to Arabic only in difficult situations. Others used simplified English rephrasing or explicitly encouraged students to speak English. Overall, participants emphasized a flexible, context-sensitive approach that balances immersion with comprehension. FT1 said, "I usually respond in English, but show understanding. If the student struggles, I rephrase the question in simple English." FT2 added, "It depends. If it is a grammar or vocabulary question, I might reply partly in Arabic; however, for casual talk, I gently remind them, 'let us try to say it in English.'" FT3 commented, "I reply in English most of the time. If I sense frustration, I use short Arabic reassurance like 'تمام، فهمت عليك' (ok I got you) then continue in English".

MT1 said he speaks in Arabic with students because, "if they know how to speak in English, they will do it, but when they speak in Arabic, I know that they need me to interact with them in Arabic also to be understandable and clear for them." MT2 reported that using Arabic depends on the situation: "If the question is about meaning or clarification, I might respond briefly in Arabic. However, if it is part of an activity or discussion, I answer in English to model the language. Sometimes I gently remind them: 'Try to say it in English.'" MT3 noted, "My approach is flexible. If the question is simple, I do my best to answer in English first and then repeat the answer in Arabic to confirm understanding. Whether a student

expects a translation depends on their individual level. Some more proficient students understand my English responses, but the majority wait for the Arabic translation to feel secure".

(d) Theme 4: Personal Teaching Beliefs and Attitudes Toward Translanguaging

Participants expressed diverse yet generally positive beliefs about the pedagogical role of Arabic. Many saw Arabic as a helpful support tool and a bridge rather than a problem. While some preferred minimizing Arabic to maintain English immersion, others believed all teachers should use Arabic strategically. Across interviews, participants consistently emphasized that Arabic is beneficial when used purposefully. FT1 stated, *"I believe Arabic can be helpful, but only as a support tool. Students feel safer and more confident when they know I understand their first language. However, my main goal is still immersion."* FT2 said, *"Arabic use is natural and realistic. It helps explain culture and reduces anxiety. I see it as a bridge, not a problem."* In contrast, FT3 noted, *"I prefer minimizing Arabic. English must surround students to think in it."* However, she added, *"But I understand sometimes we need Arabic for connection or clarification."* FT1 and FT3 reported flexible views about translanguaging, noting that they used to use *"English only"* but now *"see balance works better."* FT3 commented, *"After teaching mixed-ability classes, I now tolerate limited Arabic."* FT2 affirmed, *"I have always believed in bilingual flexibility"*.

MT2 said, *"I believe using Arabic strategically is beneficial. Complete avoidance makes students feel lost, especially at lower levels. However, overusing it makes them lazy. A balanced approach works best—Arabic for clarification, English for practice."* He noted, *"My opinion changed over time. When I was a new teacher, I tried to use only English, but I realised it was not realistic in our Saudi context. Arabic is helpful if used wisely. It supports comprehension, builds trust, and saves time. Before, I believed Arabic should never be used. "Now I see it as a tool, not an obstacle," MT3 commented, adding that using Arabic "is helpful, but strategically." Teachers should use Arabic to support learning, especially at beginner levels. The goal, however, should be a gradual shift towards using more English as the students' skills improve."* He elaborated that he changed his attitude: *"Using Arabic prevents students from feeling overwhelmed and lost in a completely English-language environment, which can demotivate them. It provides a crucial sense of security and comprehension"*.

(e) Theme 5: Participants' Perceptions of Their Comfort in Using Arabic During English Instruction

Participants reported feeling comfortable using Arabic during English instruction, noting that it feels natural and helps build trust. They highlighted that using Arabic reflects their identity, strengthening cultural alignment and allowing teachers to form emotional connections. Participants also emphasized practical advantages, explaining that Arabic simplifies explanations, saves time, and prevents instructional confusion. FT1 said, *"It feels natural and builds trust. Students see me as approachable."* She added, *"It is good to reflect on how we mix languages; it shows our teaching reality."* FT2 noted, *"I feel bilingual teaching reflects our Saudi identity; both languages belong in the classroom."* FT3 said, *"Somewhat comfortable. I prefer English, but I know Arabic helps make emotional connections with students."* However, she added, *"Translanguaging is part of who we are as bilingual educators; it deserves more recognition."*

MT1 emphasized feeling comfortable when Arabic is used. MT2 pointed out, *"I feel comfortable because it reflects my identity and connection with my students. It creates a supportive atmosphere. However, I always remind myself not to overuse it. The goal is to help students move gradually toward English independence."* MT3 commented, *"Without a doubt, using Arabic is easier for both my students and me. It simplifies explanation, saves time, and ensures that key concepts are understood without confusion"*.

V. DISCUSSION

This study investigated gendered patterns of *translanguaging* practices and attitudes among six Saudi secondary school EFL teachers through classroom observations and semi-structured interviews. The integration of classroom observation with interviews addressed an earlier limitation in existing research's reliance on self-report data (Alharbi & Alqefari, 2023). The findings reveal complex, context-dependent practices shaped by teacher gender, student proficiency, pedagogical beliefs, and institutional contexts.

Male teachers generally used Arabic more frequently than female teachers during English lessons. Among females, FT1 used Arabic the least (10%, mainly for instructions), while FT3 used it most (60%, for vocabulary and classroom management). Among males, MT1 used Arabic the least (45%, most often after asking questions), while MT3 used it most (78%, extensively for vocabulary, grammar, and classroom management).

Overall, gender influences translanguaging use in Saudi English classrooms.

As shown in Table 2, the themes and subthemes indicate that participants strategically employed translanguaging, balancing student comprehension with English immersion goals. While Arabic is frequently used to clarify content, manage the classroom, and build rapport, participants maintain clear boundaries where English-only instruction is essential. These findings provide a comprehensive understanding of how teachers navigate translanguaging as a purposeful, context-dependent practice.

Patterns of Translanguaging: Gender, Frequency, and Context

RQ1: What patterns of translanguaging between Arabic and English do male and Saudi female EFL teachers display in their classroom instruction?

The findings reveal substantial gender-related differences in translanguaging frequency. Male teachers employed Arabic more extensively (45-78% of instructional time) than female teachers (10-60%). MT3 demonstrated the highest

frequency (78%), while FT1 exhibited the lowest (10%), representing a 68-percentage-point difference. Male teachers generally employ more translanguaging, especially with low-proficiency students, to support understanding and engagement. Female teachers tend to adjust their use more according to students' grade level and English proficiency, increasing translanguaging with less proficient learners. Male teachers rely more consistently on translanguaging, while female teachers use it more selectively, reflecting different pedagogical approaches. This adaptive pattern suggests that female teachers view translanguaging as temporary scaffolding to be gradually reduced, consistent with Cenoz and Gorter's (2022) framework of flexible language boundaries that facilitate learning while maintaining target-language goals. Male teachers maintained higher translanguaging frequencies even with advanced students, suggesting they view Arabic as a legitimate, ongoing component of bilingual instruction rather than temporary scaffolding. MT3's extensive use (78%) with third-grade students challenges the assumption that translanguaging should decrease with proficiency, supporting García and Otheguy's (2020) conceptualization of drawing on students' complete linguistic repertoires.

These patterns may reflect different constructions of professional identity. Alturki and Alharbi (2024) found that Saudi female teachers frame their roles using nurturing metaphors emphasizing development and progression, whereas male teachers emphasize knowledge transmission and comprehension. Female teachers' translanguaging approach aligns with developmental scaffolding, whereas male teachers' consistent use reflects transmission-oriented approaches that prioritize immediate comprehension. This result recalls Almanea's (2025) findings on the importance of offering positive feedback to EFL learners and adds that providing feedback in L1 can be rewarding. Additionally, these results support Holmes and Schnurr's (2006) research, which demonstrates that women more frequently use language for relational purposes, while men focus on transactional communication. Female teachers' strategic modulation across proficiency levels may reflect relational approaches that adjust to student needs, whereas male teachers' consistent translanguaging reflects a transactional focus on content delivery.

However, these patterns should not be interpreted through essentialist frameworks. As Cameron (2005) emphasizes, gender operates performatively within specific contexts. In Saudi Arabia's gender-segregated system, female teachers work exclusively with female EFL learners, providing academic instruction, moral guidance, and emotional support (Al-Rasheed, 2021). Male teachers may experience different expectations emphasizing content mastery and achievement, potentially encouraging comprehensive translanguaging to ensure comprehension. Furthermore, observational data revealed additional nuances: female teachers primarily used Modern Standard Arabic (MSA), whereas male teachers frequently employed Saudi colloquial Arabic. This dialectal variation may serve different functions: MSA signals academic formality, whereas colloquial Arabic builds rapport and creates a relaxed atmosphere (Bassiouny, 2020).

Among the important findings of this study, gender is not the sole determinant of translanguaging's frequency and practices. Instead, gender differences interact substantially with contextual factors. The highest female translanguaging frequency (FT3: 60%) exceeded the lowest male frequency (MT1: 45%), indicating individual variation operates alongside gender patterns. Student proficiency emerged as the strongest predictor across both genders, supporting Macaro's (2024) evidence that strategic L1 use particularly benefits lower-proficiency learners. These findings directly address the gap in Alsagoafi's (2025) research, which documented strategic translanguaging use among Saudi secondary teachers but did not analyze gender differences despite collecting data from equal numbers of male and female teachers. The present study demonstrates that such differences are substantial and systematic, aligning with research indicating that gender shapes professional linguistic performances within specific institutional contexts (Cameron, 2005; Holmes & Schnurr, 2006).

Functions and Purposes: Gendered Pedagogical Priorities

RQ2: How do the functions and purposes of translanguaging differ between male and female Saudi EFL teachers?

Thematic analysis revealed that both genders employed translanguaging across multiple domains: teaching grammar, teaching vocabulary, classroom management, rapport-building, and supporting struggling students. Female teachers employed Arabic for contrastive analysis in grammar instruction, leveraging L1 grammatical knowledge as cognitive bridges to L2 structures and raising metalinguistic awareness. Male teachers used Arabic more extensively for direct explanation. These differences align with research suggesting that women employ scaffolding strategies that encourage independent problem-solving, whereas men favor direct information provision (Holmes & Schnurr, 2006). Regarding vocabulary, female teachers demonstrated selective translanguaging, preferentially translating abstract terms while encouraging contextual guessing for concrete vocabulary. Male teachers showed more systematic translanguaging, with one teacher reporting translating every new vocabulary item.

Female teachers employed relational management, using Arabic to build connections while redirecting behavior. Male teachers used more direct commands. Moreover, female teachers emphasized emotional connection and anxiety reduction through Arabic use in humor and encouragement. Male teachers showed variable approaches. Some used humor strategically, while others acknowledged that they do not use it for that purpose. These differences reflect gendered expectations for emotional labor in education, with female teachers expected to provide extensive relational care alongside instruction. Nevertheless, both genders extensively employed translanguaging to support shy or struggling students. This consistency suggests that supporting struggling students is a core professional responsibility transcending gender. This aligns with research demonstrating that translanguaging enhances engagement and reduces anxiety (Alsagoafi, 2025; Cenoz & Gorter, 2022).

Attitudes: Ideological Positioning and Professional Beliefs

RQ3: What attitudes do male and female Saudi EFL teachers hold toward the use of Arabic in English language classrooms?

Analysis revealed generally positive attitudes across both genders, representing a shift from traditional English-only ideologies toward acceptance of strategic Arabic use as pedagogically legitimate and culturally appropriate. These attitudes align with García and Otheguy's (2020) translanguaging theory, which views students' complete repertoires as valuable resources. Several participants reported significant attitudinal shifts over their careers. FT1 and FT3 noted they "used to use 'English only'" but "see balance works better" after teaching mixed-ability classes. MT2 reported: "*When I was a new teacher, I tried to use only English, but I realized it was not realistic in our Saudi context.*" These shifts suggest classroom experience drives pedagogical adaptation, overcoming initial monolingual ideologies (Rajendram & Thorne, 2022). MT3's evolution is particularly noteworthy: "*Using Arabic prevents students from feeling overwhelmed and lost in a completely English-language environment, which can demotivate them. It provides a crucial sense of security and comprehension.*"

Despite positive attitudes, all participants identified specific contexts where they intentionally avoided Arabic. Both genders restricted Arabic during speaking practice and productive activities. According to them, it is important to create an English environment to build fluency and confidence. This selective restriction demonstrates sophisticated pedagogical reasoning, distinguishing supportive translanguaging for comprehension from complete immersion for skill development (Cenoz & Gorter, 2022). Indeed, teachers' comfort with translanguaging was linked to cultural identity. FT2 characterized "*bilingual teaching reflects our Saudi identity—both languages belong in the classroom.*" MT2 affirmed: "*I feel comfortable because it reflects my identity and connection with my students.*" This positions translanguaging as culturally authentic rather than pedagogically compromised, challenging deficit perspectives and aligning with arguments that multilingual pedagogies must be culturally grounded (Pennycook & Makoni, 2019).

Overall, the results of the present study extend translanguaging research by demonstrating that gender operates as a significant analytical variable that should not be subsumed within homogeneous "teacher" categories, while revealing that differences are not absolute but interact with multiple contextual factors. This aligns with contemporary sociolinguistics, which emphasizes gender as performatively enacted within specific contexts (Cameron, 2005), intersecting with student proficiency, lesson content, professional identities, and institutional expectations.

VI. CONCLUSION

This qualitative study investigated gendered variations in translanguaging practices and attitudes among six Saudi secondary EFL teachers through interviews and classroom observations, addressing how teacher gender influences translanguaging in Saudi Arabia's gender-separated educational system. Results showed that male teachers employed Arabic more extensively than female teachers. Female teachers demonstrated greater responsiveness to proficiency levels, substantially increasing translanguaging with beginners while maintaining minimal use with advanced learners. Male teachers maintained consistent translanguaging across levels, viewing Arabic as an ongoing component of bilingual instruction. Nevertheless, both genders employed translanguaging across multiple domains. Female teachers used contrastive grammar instruction and selective vocabulary translanguaging; male teachers employed more extensive direct explanation in Arabic. Female teachers emphasized relational management and affective connection; male teachers used directive commands and efficiency-focused translanguaging. Both genders expressed positive attitudes toward translanguaging in EFL classrooms, reflecting shifts from English-only ideologies toward the acceptance of strategic Arabic use. However, all maintained selective English-only instruction during productive activities, demonstrating sophisticated pedagogical reasoning.

Despite its contributions, this study has some limitations that should be acknowledged. First, the small, purposefully selected sample of six secondary school EFL teachers limits the generalizability of the findings beyond similar Saudi contexts. The study was also confined to public secondary schools and did not include primary or tertiary levels, where translanguaging practices and gendered expectations may differ. In addition, although classroom observations and interviews provided rich qualitative insights, the study focused exclusively on teachers' practices and perceptions, without examining students' perspectives or learning outcomes. Future research could therefore adopt larger and more diverse samples across educational levels and institutional contexts, incorporate students' voices, and employ mixed or longitudinal designs to examine how gendered translanguaging practices evolve over time and impact language development. Further studies may also explore the interaction between teacher gender and other variables such as teaching experience, regional dialect use, or institutional language policies to develop a more comprehensive understanding of translanguaging in Saudi EFL classrooms.

APPENDIX A

https://docs.google.com/document/d/1sEDPJ1uPZo3VceHSrXr_4pyJhVoSzme4/edit

APPENDIX B

<https://docs.google.com/document/d/1INXVQ9lyjJL3A10KXtOxov18RQFo0qdi/edit>

APPENDIX C

<https://docs.google.com/document/d/10Kn6ixTZLfw6ZleQAAZ0eIT8l9AJBu3k/edit?usp=sharing&ouid=105201631064797173537&rtpof=true&sd=true>

ACKNOWLEDGEMENTS

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (Grant Number IMSIU-DDRSP2603).

REFERENCES

- [1] Alasmari, M., Qasem, F., Ahmed, R., & Alrayes, M. (2022). Bilingual teachers' translanguaging practices and ideologies in online classrooms in Saudi Arabia. *Heliyon*, 8(9), e10537. <https://doi.org/10.1016/j.heliyon.2022.e10537>
- [2] Almanea, M. (2025). Less is NOT more for learners: EFL learners' preferences and perceptions of teachers' written corrective feedback. *Acta Psychol.*, 255, 104926. Doi: 10.1016/j.actpsy.2024.104926
- [3] Almanea, M., Mohsen, M., & Mahdi, H. S. (2024). Is “The Earlier, The Better” for EFL in Saudi Arabia? Insights from Early Grades EFL Teachers at Saudi Primary Schools. *International Journal of Learning, Teaching and Educational Research*, 23(10), 123–155.
- [4] Al-Rasheed, M. (2010). *A history of Saudi Arabia* (2nd ed.). Cambridge University Press.
- [5] Alshammari, M. M. (2011). The use of the mother tongue in Saudi EFL classrooms. *Journal of International Education Research (JIER)*, 7(4), 95-102. <https://doi.org/10.19030/jier.v7i4.6055>
- [6] Alharbi, M., & Alqefari, A. (2023). Attitudinal stance toward translanguaging in Saudi Arabia: A mixed-method case study. *International Journal of Applied Linguistics*, 33(1), 47–65. <https://doi.org/10.1111/ijal.12448>
- [7] Alqahtani, M. H. (2022). The Saudi 2030 vision and translanguaging in language learning in Saudi Arabia: Looking for concord in the future. *Journal of Language and Linguistic Studies*, 18(Special Issue 1), 556-568.
- [8] Alsagoafi, A. A. (2025). Pedagogical translanguaging in practice: Creating spaces by softening the boundaries between languages. *Journal of Language Teaching and Research*, 16(6), 2090-2097. <https://doi.org/10.17507/jltr.1606.31>
- [9] Alturki, B. M., & Alharbi, M. A. (2024). Saudi Female English Teachers' Professional Identity Construction: A Qualitative Study of the Internal and External Factors. *Forum for Linguistic Studies*, 6(3), 509–528. <https://doi.org/10.30564/fls.v6i3.6529>
- [10] Al-Wer, E., Horesh, U., Herin, B., & De Jong, R. (2022). *Arabic sociolinguistics*. Cambridge University Press. <https://doi.org/10.1017/9781316863060>
- [11] Bassiouney, R. (2020). *Arabic sociolinguistics: Topics in diglossia, gender, identity, and politics* (2nd ed.). Georgetown University Press.
- [12] Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- [13] Cameron, D. (2005). Language, gender, and sexuality: Current issues and new directions. *Applied Linguistics*, 26(4), 482–502. <https://doi.org/10.1093/applin/ami027>
- [14] Canagarajah, S. (2011). Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review*, 2, 1–28. <https://doi.org/10.1515/9783110239331.1>
- [15] Cenoz, J., & Gorter, D. (2022). Pedagogical translanguaging and its application to language classes. *RELC Journal*, 53(2), 342-354. <https://doi.org/10.1177/00336882221082751>
- [16] Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.
- [17] Conteh, J., & Brock, C. (2021). *Safe spaces and learning places: Translanguaging for equity in multilingual education*. Multilingual Matters. <https://doi.org/10.21832/9781800414066>
- [18] Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage Publications.
- [19] Einarsson, C., & Granström, K. (2002). Gender-biased interaction in the classroom :The influence of gender and age in the relationship between teacher and pupil. *Scandinavian Journal of Educational Research*, 46(2), 117-127. <https://doi.org/10.1080/00313830220142155>
- [20] Elyas, T., & Badawood, O. (2016). English language educational policy in Saudi Arabia post-21st century: Enacted curriculum, identity, modernization. *FIRE: Forum for International Research in Education*, 3(3), 70-81. <https://doi.org/10.18275/fire201603031093>
- [21] García, O., Johnson, S. I., & Seltzer, K. (2017). *The translanguaging classroom: Leveraging student bilingualism for learning*. Caslon.
- [22] García, O., & Otheguy, R. (2020). Plurilingualism and translanguaging: Commonalities and divergences. *International Journal of Bilingual Education and Bilingualism*, 23(1), 17-35. <https://doi.org/10.1080/13670050.2019.1598932>
- [23] Garcia, O., & Wei, L. (2014). *Translanguaging: Language, Bilingualism, and Education*. New York, NY: Palgrave MacMillan. <https://doi.org/10.1057/9781137385765>
- [24] Holmes, J., & Schnurr, S. (2006). *Gendered talk at work: Constructing gender identity through workplace discourse*. Blackwell.
- [25] Li W. (2018). Translanguaging as a practical theory of language. *Applied Linguistics*, 39(1), 9–30. <https://doi.org/10.1093/applin/amx039>
- [26] Lin, A. M. Y., & He, P. (2022). Translanguaging as translanguaging practice: A critical review and reconceptualization. *Critical Inquiry in Language Studies*, 19(1), 1–28.
- [27] Litosseliti, L. (2006). *Gender and language: Theory and practice*. Routledge.

- [28] Macaro, E. (2024). First Language Use and Language Learner Strategies in EMI Higher Education. In A. B. M. Tsui & E. Macaro (Eds.), *Language Issues in English Medium Instruction: Theoretical Orientations and Cases from Disciplinary Practitioners* (pp. 14-28). Routledge. <https://doi.org/10.4324/9781003476313>
- [29] Myers-Scotton, C. (1993). *Social motivations for codeswitching: Evidence from Africa*. Oxford University Press.
- [30] Paulsrud, B., Rosén, J., Straszer, B., & Wedin, Å. (2017). *New perspectives on translanguaging and education*. Multilingual Matters.
- [31] Pennycook, A., & Makoni, S. (2019). *Innovations and challenges in applied linguistics*. Routledge.
- [32] Rajendram, S., & Thorne, S. L. (2022). Teacher education for translanguaging pedagogy: Shifting ideologies and practices. *TESOL Quarterly*, 56(2), 634–661.
- [33] Sadker, D., Sadker, M., & Zittleman, K. R. (2009). *Still failing at fairness: How gender bias cheats girls and boys in school and what we can do about it*. Scribner.
- [34] Turnbull, M., & Dailey-O'Cain, J. (Eds.). (2009). *First language use in second and foreign language learning*. Multilingual Matters.
- [35] Vision 2030. (2021). *Human Capability Development Program. Kingdom of Saudi Arabia*. Retrieved January 18, 2026, from <https://www.vision2030.gov.sa/en/explore/programs/human-capability-development-program>
- [36] Vogel, S., & García, O. (2021). Translanguaging. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Wiley. <https://doi.org/10.1002/9781405198431.wbeal1488.pub2>

Sita Naji Aldossari is currently pursuing a Ph.D. in Applied Linguistics at Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia. She holds a Master's degree in Teaching English for Speakers of other Languages from King Saud University in Riyadh, Saudi Arabia, and a BA in English Language from Princess Nourah bint Abdulrahman University, Saudi Arabia. With over a decade of experience teaching English in Saudi Arabia, her expertise lies in secondary and Intermediate grades education, Early childhood education, and Training teachers in several fields. Her research interests in online learning, English language teaching, students autonomy and motivation and integrating Artificial Intelligence in teaching English language. (ORCID iD: 0009-0000-9786-5278) Email: 446011952@sm.imamu.edu.sa

Manar Almanea received the B.A. (Hons.) and M.A. degrees in applied linguistics from Princess Norah Bint Abdulrahman University and the Ph.D. degree (with Distinction) in applied linguistics from King Saud University. She is currently a professor of applied linguistics with the College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia.

Her research interests include linguistics and EFL learning and teaching. Her works are published in top-tier journals like *Computer Assisted Language Learning*, *Acta Psychologica*, *Psycholinguistics*, *IEEE Access*, among others. She has earned several scientific awards and is a member of numerous scientific and linguistic societies, both nationally and internationally. <https://orcid.org/my-orcid?orcid=0000-0002-3615-2394>

Sisterhood in Leila Aboulela's *Bird Summons* and Toni Morrison's *Sula*

Nadera Nader Al-Staity
Middle East University, Amman, Jordan

Nasaybah Walid Awajan
Middle East University, Amman, Jordan

Sahar Mohammad Tarawneh
Middle East University, Amman, Jordan

Abstract—This study aims to explore how Toni Morrison and Leila Aboulela represent the importance of sisterhood in the lives of women in diaspora, in their novels, *Sula* (1973) and *Bird Summons* (2019). The study attempts to highlight the challenges that women in diaspora face, such as marginalization, patriarchy, and discrimination. Therefore, sisterhood has emerged as a powerful force for empowerment and resistance to colonial contexts. These two novels are analyzed through the lens of Postcolonialism and Feminism. Hence, the researchers rely on the views of Edward Said, Franz Fanon, Gayatri Spivak, and Bell Hook. The study concludes that both Morrison and Aboulela in their novels explore the significant role of sisterhood in the lives of these women in offering solidarity and mutual and collective empowerment in colonial contexts, particularly to those who share common experiences and challenges. The contribution of the study lies in connecting the two novels and investigating how both writers share the same concept of sisterhood.

Index Terms—sisterhood, *Sula*, Morrison, *Bird Summons*, Aboulela

I. INTRODUCTION

The concept of sisterhood has been influential in feminist movements and women's rights advocacy, as it promotes the idea that women can achieve greater life quality and overcome systemic barriers by standing together in solidarity. Additionally, sisterhood can be a source of emotional support, encouragement, and mentorship for women in various aspects of their lives. Overall, sisterhood celebrates the unique bond between women and highlights the importance of building and maintaining positive, supportive relationships among women in personal, social, and professional spheres. Bauder and Juffs (2019) state that sisterhood considers the continuous efforts of women to reach the level of equality they dream of, and their tendency to believe that they are equal in terms of obtaining their rights. The concept of sisterhood in Afro- American literature was built by highlighting it in all kinds of women's literature, and its repetition in literary works, as in the works of Morrison and most of the stories of black women, where writers and critics call for women unity to gain strength and prove their ability to resist oppression.

Parashar (2016) suggests that the Feminist movement did not develop and become in its current form at once. Rather, its history is divided into three stages. The first stage, which started in the late nineteenth century, focused on the rights of women. The second stage began in the early sixties of the twentieth century and at this stage, the scope of the dialogue expanded to include other issues such as the workplace and issues of sexuality and highlighting the struggles that Afro-Americans face due to slavery, racial discrimination, and patriarchy. Sachdev (2020) explains that in Afro-American literature, women's works carry the mission of building sisterhood within the feminist alliance, and their works are important to achieve social justice for black women. Afro-American literature played an important and pivotal role in the history of women's literature and in highlighting the importance of sisterhood among women, where this literature sheds light on the corruption of the traditional image of the silence and subjugation of women, and presents a strong image of the educated and assertive woman (Yousef, 2023, 2024; Abuhassirah & Salameh, 2024).

Toni Morrison's writings focused on women's experiences in black American communities and the importance of sisterhood in their lives (Wagner, 2022). This theme of sisterhood is also reflected in diasporic literature. Smith (2021) states that in diasporic literature, people experience adapting to new countries or new languages, cultures, and societies. The writers of this type of literature focus on highlighting the experiences of women in diaspora, because of the gap they fall into between their traditions and what they need to do to integrate into modern society. In addition to themes of love, loss, and identity, diasporic literature includes stories of immigrant women, the reflection of sisterhood, and the issues they encounter in their homes or the host land. This genre has given birth to icons such as Leila Aboulela, Mohja Kahf, Anita Desai, and others.

Aboulela, a diasporic writer, is an Arab Sudanese, who lives in England. In her works, she focuses on the independence of Arab women, and that their roles are not confined to home only (Omet, 2022). Furthermore, sisterhood emerges as a

form of resistance in the face of postcolonial implications that are imposed on women. Women in diaspora suffer from stereotypical images that were imposed on them according to the Western views of the East. Edward Said in *Orientalism* 1978 discusses the representation of the East through the orientalist discourse, how the West misrepresent the East by labeling them as being barbaric, uneducated, and backward, to justify the Western hegemony and the intervention in the Eastern societies. Therefore, as a result, oriental women are seen as passive, submissive, and sexual objects.

II. METHOD

The researchers use postcolonialism and feminism to approach the two novels, to discuss the theme of sisterhood. The researchers rely on the views of Said, Fanon, Spivak's, and Hooks. Postcolonialism is the period that comes after the end of colonialism, which cares about exposing the effect of colonialism on people who have undergone colonial experience. Postcolonialism highlights the predicaments that the colonized world went through during Eurocentric colonial hegemony. Many postcolonial writers started to write back to change the stereotypical images that were created by the West during the European colonial aggression (Childs & William, 1997).

Feminism is a literary theory that calls for women's rights, and their continuous attempt to fight oppression, which paved the way for their liberation, feminism as a movement calls for enhancing the life quality for women and engaging them in the political, social, economic fields (Arizah, 2020). In viewing the Subaltern discourse, Spivak highlights the notion of women's ranking in society, and she integrates Postcolonialism with feminism where she connects the issues that women face with their ranking in society, which is determined by men. Spivak examines women's relationship with men as the relation between the colonized and the colonizer, where women are looked down at, and considered as unequal to men. Moreover, they are marginalized and positioned as the "other" (Rose, 2014).

III. DISCUSSION

The Representation of Sisterhood in Morrison's *Sula*

The story of *Sula* is set in the fictional town of Medallion, Ohio. The novel primarily focuses on the lives of two Afro-American women, in a black community Sula and Nel. The Narrative begins in the early 1900s, when slave descendants inhabit a place up on the hill, which they call "The Bottom" an ironic name that represents the condition of the African American people. In the Bottom, Nel Wright and Sula Peace find each other as young girls, and they become close friends. Sula comes from a chaotic household and is seen as somewhat unconventional, while Nel comes from a more traditional and conservative family. Despite their differences, they develop a deep bond and become inseparable. Nel Wright and Sula Peace's personalities are shaped according to their surroundings and the way they are raised in their households. To begin with Nel's relationship with her mother. Nel's mother Helene is a domineering, conservative woman. Helene's mother was a prostitute for it was her occupation. Therefore, Helene's grandmother religiously raised her, in order to avoid following her mother's example (Morrison, 1973).

On the other hand, Helene's husband is always absent, and she finds herself alone in running her household, and raising her child. Therefore, Helene follows her grandmother's footsteps but with more exaggeration. Helene smothers her daughter while raising her, and she succeeds in molding Nel's personality the way she wants, submissive, obedient, and silenced to fit in with the black community ideals, to be, a well-mannered, proper woman, whose only dream, is to find a good husband (Morrison, 1973).

Moreover, Nel's heartless mother does not only want to determine her daughter's fate but also wants to change her physical appearance. Helene is not fond of Nel's "broad flat nose (although Helene expected to improve it somewhat)" (Morrison, 1973. p. 18). Hence, Helene forces Nel to put a clothespin on her nose. "go 'head and pull your nose. It hurts Mamma. Don't you want a nice nose when you grow up?" (Morrison, 1973, p. 55).

In this perspective, living in a patriarchal society in the time of slavery, and after the emancipation proclamation, Afro-American women find themselves in an extremely sensitive position. Pascual (2020) argues that society is divided into three sections, on the top of the social status based on race and sex, are white men, then followed by white women and black men, and finally, at the bottom are black women, and in some cases, they are not categorized. Therefore, black women suffered tremendously for being poorly treated by the white community, and the black male.

In this sense, Spivak (2009) points out that the Subaltern as a concept means; marginalized, inferior people who are suffering under colonization. Spivak as a postcolonial writer does not only speak out for the colonized people in the third world, but she also addresses specifically postcolonial women around the world. Moreover, she encourages them to speak up against injustice and social oppression. In their paper, Ouakouk and Bekkari (2023) argue that "Spivak's work she speaks heavily of the injustice and harmful oppression against women in general, and Third World women and non-Europeans in specific. She expressed her desire to deliver and possibly give a voice to the subalterns who cannot speak or are silent" (p. 9). Ouakouk and Bekkari (2023) elaborate further, that Subaltern women are subjected to all sorts of injustice, such as verbal and physical harassment, discrimination, rape, sexism, and violence.

Although *Sula* is set after the emancipation proclamation, Morrison sheds light on the status of Afro-Americans in the aftermath of slavery. Washington (2012) argues that although the slavery period has ended, Afro-Americans should be treated as equal to white people. Helene faces this racist segregation when she is summoned to her grandmother's funeral, her grandmother's house is in the city, Therefore, she goes with little Nel by train, however, mistakenly, she enters the

White cabin and does not notice the marked door for “colored only”. This mistake puts her in an embarrassing situation with the white conductor, who addresses her aggressively and asks her to go back to the colored cabin (Morrison, 1973). Moreover, Helene faces another racist incident, when she and her daughter want to use the bathroom, but since black people are prohibited from using the public facilities with whites. Helene and her daughter squat on the ground hiding between the grass by the train along with the black women and their children, while white people are watching them from the train windows (Morrison, 1973).

Since Afro-American women are at the bottom of the hierarchical pyramid. They are struggling in a racist sexist patriarchal society. They struggle to find a decent job to support their families, after being abandoned by their men who they focus solely on how to maintain their social status. Therefore, they find themselves the only providers in their families. Pascual (2020) also explains the absence of black men from the familial picture by saying that “Most of them are pictured as almost invisible figures within their household, that they either abandon their families or betray their wives and abandon the household” (p. 30). Therefore, Afro-American men do not only protect and support their women from racial aggression, but they encourage the dehumanization of African American women. As Hook (1981) states, “Most black slaves stood quietly as white masters sexually assaulted and brutalized black women and were not compelled to act as protectors” (p. 35). As a result, for black women to support their families, some of them work as concubines. However, this act puts their reputation at risk. Afro-American women are not only seen as inferior, and marginalized, but as the embodiment of female evil and sexual lust. Hooks (1981) states that “Black women were naturally seen as the embodiment of female evil and sexual lust. They were labeled Jezebels” (p. 31).

Nel’s personality is shaped according to her mother and her surroundings. Nel at a young age witnesses remarkable racist incidents; therefore, she knows her current status quo. In addition, her mother’s obsession with forcing little Nel to fit into a stereotypical culture is determined by a patriarchal community. Nel, as a result, becomes submissive, and vulnerable, seeking affection and support.

After the death of her father, Sula leaves with her mother Hanna, to her grandmother’s house Eva. When Eva’s husband Boy abandons her, she has three children to feed. Therefore, she leaves her children with the neighboring family to secure a job, and after eighteen months, she comes again for her children with one leg and money. People in Eva’s community “said Eva stuck it under a train and made them pay off. Another said she sold it to a hospital for 10.000\$” (Morrison, 1973, p. 31). Eva, with this money, builds a big house with many rooms and floors, and she opens it to anyone who needs shelter.

During her stay at her grandmother’s house, Sula is detached and secluded. Her mother is not taking care of her nor giving her love and affection. Hana is always busy making relationships with as many men as she can., Sula always wonders if her mother loves her, however, on one occasion she hears her mother telling her friends, “I love Sula, but I don’t like her” (Morrison, 1973, p. 57). Eva’s house misses law and order and with the absence of parental guidance, Sula becomes hostile, carefree, and rebellious.

Even though Nel and Sula are the opposite of each other in terms of their lifestyles and characters. However, they share the same struggle and discomfort. Therefore, when they met, they knew they completed each other, and with their bond, they supported and took care of each other (Morrison, 1973). Therefore, through their bond, they skip what bothers them, and focus solely on themselves, not only to empower themselves but more importantly, to heal their spiritual wounds.

Nel and Sula begin to discover how their bond can empower them unconsciously. Further, how this bond is going to affect their personalities accordingly. As for Nel, as I mentioned earlier, due to her relationship with her domineering mother she lacks confidence and self-respect. However, Nel with Sula feels comfortable in her skin and grows confident. Hence, she stops caring about her physical appearance and focuses more on her mind and soul. Nel after she meets Sula, accepts the natural beauty that distinguishes her, and learns how to embrace it. Nel stops using her nose pin and she “slid it under the blanket as soon as she got in the bed” (Morrison, 1973, p. 55). Moreover, she stops worrying about “the hateful hot comb” that she has “to suffer through each Saturday evening” (Morrison, 1973, p. 55) when her mother straightens her hair every Saturday evening.

As for Sula, as mentioned before, Sula has an agitated temper and a chaotic mind, she lacks a sense of belonging. Therefore, she always feels detached. With Nel, Sula feels cared for, her existence now matters to someone. Besides, when Sula is upset about the way her mother treats her, Nel comes to her “pulling her away from dark thoughts back into the bright, hot daylight” (Morrison, 1973, p. 57). Moreover, Sula does not understand how to show affection, and she “could hardly be counted on to sustain any emotions for more than three minutes” (Morrison, 1973, p. 53). Sula now is calmer and self-composed in the presence of Nel.

Moreover, Nel and Sula enjoy each other’s company and their souls become so intimate to the extent that they display their unity and bond through their little game. This is shown in an incident when Nel and Sula as young girls start playing in the grass play, they dig two holes, where “together they worked until the two holes were the same” (Morrison, 1973, p. 58).

Weems (2004) argues that Afro-American women have always shared an unbreakable bond where each member gives and receives equally. They form a supportive community demonstrating a strong sense of responsibility for one another and showing empathy for shared experiences. Their interactions are based on love, including constructive criticism, and the exchange of common and individual experiences. Nel and Sula know as black girls that through sisterhood they support and empower each other.

Furthermore, Nel and Sula as black girls, are learning that through sisterhood, they can love themselves and each other, while this love can be implemented in different forms such as care, support, empowerment, and the ability to stand for each other in a time of need. Furthermore, through sisterhood, they created a sense of selflessness as a result; they established an efficient relationship based on mutual nurturing, self-love, and the ability to erase insecurities. Oliveira (2011) points out that “Sisterhood creates the possibility for black women to develop their subjectivities by aligning themselves with girlfriends” (p. 50).

Moreover, Morrison demonstrates through *Sula* that sisterhood means making sacrifices and being selfless when it comes to standing up for each other. In this scene, Sula feels responsible for protecting Nel, when some boys try to harass Nel on her way to school, Sula rushes to protect Nel, and immediately, in an irresponsible act, she cuts her finger with a knife to scare off the boys, and eventually she succeeds (Morrison, 1973). Beckerman (1981) describes Sula’s act as irresponsible. From another perspective, Morrison tries to intrigue through *Sula* the reader into questioning the validity of sisterhood, when this bond faces life challenges.

Furthermore, Morrison puts Nel and Sula in two incidents, which jeopardize their relationship, and how they react accordingly. The first incident occurs when Nel and Sula were young and playing by the river. They meet a little boy named Little Chicken. Nel and Sula offer to play with him, and Sula “picked him up by his hand and swung him out [...] When he [Little Chicken] slipped from her hands and sailed away out over the water” (Morrison, 1973, pp. 60-61). Assumed to be dead, Nel and Sula are consumed with guilt thinking that they are responsible for drowning Little Chicken. As a result, Sula starts crying and collapses, but Nel is beside her comforting her and telling her “It ain’t your fault” (Morrison, 1973, p. 63). Sula feels safe and “stood up and allowed Nel to lead her away” (Morrison, 1973, p. 63). At the time of the investigation, Nel and Sula “did not touch hands or look at each other, there was a space, a separation between them” (Morrison, 1973, p. 64). Sula thinks that at any moment the sheriff will point at her and accuse her of murdering.

The other incident is when Nel and Sula grow up, Jude Greece a young ambitious black man, plans to marry Nel. Jude dreams of marrying someone who does not treat him like his mother, someone who can share with her his dreams, hopes, and plans. Jude wants to marry someone “who had never seemed hell-bent on marrying, who made the whole venture seem like his idea, his conquest” (Morrison, 1973, p. 83). This indicates the mentality of black men in how to be superior and domineering, forcing their desires and needs upon black women, Jude as the other black men desperately want to change their status in their community and climb the hierarchical ladder, and always unsatisfied with their status quo. However, Nel’s personality is quite the opposite. Morrison points out that Nel is a different person from Sula, Nel feels comfortable with Sula, and for Nel, Sula is her haven to go to if someone or something bothers her (Morrison, 1973).

Eventually, Nel chooses to be a domestic wife. However, this choice costs her to accommodate Sula’s friendship (Morrison, 1973). Nel withdraws from Sula’s life to prepare herself for her new life. Dharitri (2019) illustrates that “The two friends are compelled to start a life of their own by the society” (p. 3). Nel decided to marry according to the conventions of society. This shows how society-controlled women in the society through different social issues like marriage. Sula on the other hand chooses to leave, when she sees that Nel is no longer close to her, Nel now is busy in her new life (Morrison, 1973). Therefore, Sula leaves the town to experiment with something new, away from the societal conventions.

In this perspective, Morrison demonstrates to the reader the consequences of not maintaining sisterhood through Nel and Sula. After ten years, Sula returns to her hometown, and she becomes a replica of her mother, not different from Sula, Nel also becomes exactly like her mother. Sula is rebelling against societal conventions, she is having many affairs with men as her mother was, and she is careless and ruthless. On the other hand, Nel becomes a mother of three children always visiting the church and doing her sermons, she makes sure that everyone in town knows that she is an obedient housewife and follows the society norms. Sula visits Nel at her house and sees how Nel has changed and has three children, she is now occupied and there is no room for Sula in her life.

Nonetheless, Nel and Sula choose different paths. Sula chooses to be rebellious and does not abide by society’s rules, while Nel chooses to be obedient, and follows society’s instructions. However, none of them feels as happy or comfortable as when they were together. Sula feels deserted and outcasted, she does not differentiate between what is right and what is wrong, and she lives her life carelessly. While Nel yearns for a sense of belonging, she misses how to be loved by someone, someone who can understand her and feel comfortable with her, although Nel thinks that her husband Jude will compensate for the loss of Sula. However, she was wrong, Jude like any other Black man thinks of himself and his desires and focuses on pursuing his dreams alongside white people.

Nel and Sula’s reunion makes it possible for their bond to be established again because they need each other. However, everything falls apart, when Sula makes love with Jude, Nel’s husband, Nel feels betrayed, and cannot fathom the idea that her only, and dear friend could betray her. This second incident affects Nel and Sula’s bond forever, Nel could not forgive Sula for her infidelity with her husband. On the other hand, Sula with her troubled life and mind could not understand Nel’s reaction, considering the nature of their relationship she thinks, “She had no thought at all of causing Nel pain when she bedded Jude” (Morrison, 1973, p. 119). Sula also thinks that she never understood the concept of marriage for “having lived in a house with a woman who thought all men available” (Morrison, 1973, p. 119).

From the previous discussion, Nel and Sula’s behavior refers to the historical and social aggression that Afro-American women witnessed. Miller (2023) argues that the psychological effect of slavery and trauma on Afro-American women have been profound and long-lasting, the experiences of slavery including dehumanization, and violence, have had a

significant psychological impact on Afro-American women and their descendants. The historical trauma has become embedded in the collective memory of the Afro-American community. Afro-American women continue to experience racial trauma because of systematic racism, discrimination, and microaggression, contributing to ongoing mental health challenges.

This discussion leads us to Fanon's violence theory. Hilton (2014) examines Fanon's theory of the psychological effect of colonization and oppression. He argues that the experience of being colonized leads to profound psychological trauma and a sense of inferiority in the colonized people. Therefore, Fanon views the use of violence by the colonized as a response to the structural violence and dehumanization inflicted upon them. As a result, Sula rejects an obedient subjugated black woman, and she wants to live her life by her rules.

Shovon (2022) argues that "The main cause of the damage to the friendship is their "conflicting modes of moral perception" Sula learns to take care of herself and takes responsibility for her actions in ten years and thus becomes confident about her activities. In these years, Nel learns to take care of her husband and her children and becomes a little oblivious about herself. She feels insecure and loses her confidence because, in these years, she forgets to express herself" (p. 3).

Nel and Sula after this incident separate again. Nel believes that her life has been ruined because of Sula, therefore, she spends the time hating Sula and regretting her relationship with her. Nel misses Jude and thinks that he stopped since she left. Nel could move on with her life and hence fall into despair.

One day Nel hears that Sula is sick, and she cannot prevent herself from visiting her. In Sula's room, Nel sits and confronts her, and starts arguing about the reason that made Sula betray her only friend. While Sula is focusing on their friendship, their bond is stronger to be ruined by this insignificant incident. Sula reminds us that "if we were good friends, how come you couldn't get over it" (Morrison, 1973, p. 145). Nel and Sula in their subconscious are certain that they need each other more than they need men because they know how men are in their society. However, Nel is driven by her anger and denies the truth.

After a while, Sula dies because of her illness alone. Nel visits her grave in the cemetery, remembering every little detail about Sula, while Nel is walking, she stops "Her eyes twitched and burned a little. Sula? She whispered, gazing at the top trees. Sula?" (Morrison, 1973, p. 174). Nel now understands Sula's words, she understands that she needs Sula more than anyone in the world. Nel's final words are "All that time, all that time, I thought I was missing Jude." And the loss pressed down on her chest and came up into her throat. "We were girls together," she said as though explaining something. "O Lord, Sula," she cried, "girl, girl, girl girl girl." It was a fine cry – loud and long- but it had no bottom and it had no top, just circles and circles of sorrows" (Morrison, 1973, p. 174).

The Representation of Sisterhood in Aboulela's *Bird Summons*

In this section, the three main characters Salma, Moni, and Iman are discussed concerning sisterhood. The first character to discuss is Salma, who is originally from Egypt, her age is around forty. She is married to a Scottish man, and she has four children, she works as a massage therapist. Salma has a strong leading personality, when she speaks "people listen" (Aboulela, 2019, p. 16). Salma is a successful community housewife, she has a good job, and a good relationship with her husband who loves her and gives her freedom to do whatever she wants, moreover, her "four children were burly and good at everything: school, sports, hobbies" (Aboulela, 2019, p. 6). However, she could not stop the feeling of the cultural displacement "she [Salma] was the one who must always be making the effort to belong. Digging deeper all the time, craving connections, self-conscious that her roots, despite the children, might not be strong enough" (Aboulela, 2019, p. 33). Salma feels a deep longing for a connection to her home culture and struggles with the challenges of assimilating into a new society.

Back in Egypt, Salma first studied medicine, but her degree was not accepted in Britain, because it "was not sufficient" (Aboulela, 2019, p. 9), therefore, she tried to do the qualifying test, but she failed twice. Salma, heartbroken, starts to find other alternatives and trains herself to become a massage therapist. Moreover, Salma always feels alienated and distant from her family. She believes that her kids have a special language with their father, David. Salma understands every word they say, yet she does not know what they are talking about, she "would then feel that they were his [David] and not hers. She was the outsider, the foreign wife, and they were one unit" (Aboulela, 2019, p. 33).

Moreover, Salma's identity is fragmented, as she navigates between her Egyptian background and her life in Britain, Salma remembers how was to live in Egypt and misses life there, how she would take the bus with Amir, heading to have "Foul and Tamiyah and tea with mint" (Aboulela, 2019, p. 34) to eat for breakfast. Salma and Amir were to be engaged, however, everything changed when she traveled to Britain.

The second character is Moni, who is a Sudanese woman, who has a disabled son Adam who has severe cerebral palsy. Moni dedicates her life to taking care of her son neglecting herself and her husband Murtada, Moni rejects her husband's proposal to go with him as he has a good job offer in Saudi Arabia. Moni prioritizes her son's well-being and the medical care he receives in Britain.

Furthermore, Moni accuses her husband of being indifferent about his son's situation, while Murtada begs her to move on and return to her old self, how Moni and before Adam was positive and successful "with her higher-powered bank job and independence" (Aboulela, 2019, p. 13). Moni's struggle revolves around changing societal norms and stereotypes regarding disability and motherhood. Moni wants to find her own identity and break free from the confines of societal expectations. In Sudan when Moni and Murtada went to visit his family, for his brother's wedding, people were cruel and

judging about Adam's health status (Aboulela, 2019). Therefore, she would keep Adam at her room all the time to keep away predators off him and her and promised herself that she would never go to Sudan because people there do not understand.

The third character is Iman, who is a Syrian civil war refugee, and she comes to Britain seeking for better life. Iman longs for stability in her life as she is "in her twenties but on her third marriage. Once widowed, and once divorced" (Aboulela, 2019, p. 3), moreover, she yearns to be reunited with her scattered family members, especially her mother. Iman's second marriage allowed her to go to Britain but unfortunately, she got divorced after her husband was imprisoned. Iman meets Ibrahim in Britain, Ibrahim is a student at the university away from his family and alone in Britain, he is always afraid to have affairs with women, therefore, to prevent himself from sins, he seeks marriage, when he sees Iman, he is infatuated with her beauty, and marries her secretly, without the knowledge of his parents.

Furthermore, Iman's character is marked by her obedience and dependence on others. She consistently follows instructions and feels powerless to alter her destiny, despite her strong desire for a new sense of identity and self-awareness.

It is worth mentioning that all three characters are subjected to the stereotypical image that haunts Arab Muslim immigrants in Europe. Said examines the relationship between the East and the West, he argues that the Western knowledge about the East, particularly the Arab and Islamic world, is not based on objective understanding but rather on a set of preconceived notions and cultural biases. Said (1978) describes the Western academic and cultural discourse that perpetuates these stereotypes, portraying the Orient as exotic, backward, and inferior.

Aboulela sheds light on this issue, specifically Arab Muslim women who are misrepresented and marginalized in Western societies. Arab Muslim women in diaspora often find themselves at the intersection of multiple stereotypes and cultural expectations. They are not only subjected to orientalist depictions of being exotic and submissive Middle Eastern women but also face Islamophobia and the assumption that they are oppressed because of their religious and cultural background.

Salma, Moni, and Iman's unexpected friendship transcends their differences. Their interactions and shared experiences highlight the power of sisterhood. Despite their struggles and conflicts, they find strength in their collective support for one another.

Salma meets Iman throughout the Arab Muslim women group, that Salma leads. Since then, Iman has always been with Salma. Iman considers Salma as her family, for example, when Iman marries Ibrahim "Salma had helped her with all the preparations, hosting the henna party at her house and even sugaring Iman's legs for her" (Aboulela, 2019, p. 10).

On the other hand, Salma thinks of Iman as if she is her daughter, when the gap between Salma and her children widens as they grow older and "they became more British and less piece of her" (Aboulela, 2019, p. 7), she comes closer to Iman, Iman is always there when Salma needs her. Iman always knows Salma's intentions; therefore, she does not feel the need to justify her actions all the time.

Moni and Salma meet when Moni takes Adam for his massage therapy session, however, she does not find the two British therapists that she is accustomed to, rather, she finds Salma taking Adam to the session. However, Moni questions Salma's abilities, for being an Arab descendant, which indicates how Arabs are looked at and compared to the Westerners. Moni thinks that "Salma would neither be as professional nor as qualified as her British counterparts" (Aboulela, 2019, p. 30). However, Moni is surprised, when she finds that Salma "was even better with Adam" (Aboulela, 2019, p. 30) than the other British therapists. Moreover, Salma offered to give a session to Adam "free of charge" (Aboulela, 2019, p. 30) during Christmas vacation when the center was closed.

Since then, Salma and Moni have become friends. Moni admires Salma for the support she has shown amid her misery, and she feels secure knowing she has someone kind and supportive by her side. Also, Moni benefits from Salma as she works in a medical care institution and has all the medical knowledge and the proper practice, Moni can ask about Adam's health condition when it is needed. On their way to their spiritual journey Salma, Moni, and Iman start questioning the role of patriarchy in their lives, and what to expect from men within the diasporic context. Tom (2021) argues that women in diaspora may find themselves caught between the cultural expectations and their ethnic heritage and the opportunities for gender equality in the host culture. She adds that sisterhood provides a safe and supportive space for women to discuss also to challenge patriarchal norms, and to empower each other.

As for Moni, she believes that her husband is not supportive and does not care about his son's health condition. Moni feels alone struggling with her disabled son, her all she wants from her husband is to be by her side, reassuring her that Adam is going to be better. Moreover, Moni feels that she carries a huge burden on her shoulders, and she wishes that Murtada would take the lead so she could rest. However, Murtada's only concern is to be in a place where he can be himself in an Arabic country, and find a decent job, a place where he cannot be judged or always have to feel the pressure to live according to the West lifestyle (Aboulela, 2019).

As for Salma, Salma always compares her husband David with her Egyptian friend Amir. Salma admires her husband and admits that he loves her, he is supportive of her and also helps her in the house and the kids. However, she feels that he is indifferent when it comes to her wishes and needs, for example, when Salma's medical degree was rejected in Britain and failed twice in the qualifying exam, David did not encourage her to pursue her dream of becoming a doctor, on the other hand, Amir was always encouraging her to follow her dreams. Amir always addresses her as Dr. Salma (Aboulela, 2019), as he knows this is the only thing that she ever wanted in her life. Therefore, Salma, after all these years cannot forget Amir and starts to contact him over the phone, Salma feels alive again, she believes that Amir knows what she

wants, is capable of feeding her ego, boosting her self-esteem, and acknowledging her capabilities.

Finally, Iman questions the difference between marriage and prostitution, with her unfortunate marriages, she loses track. Iman was fifteen years old when she was married, therefore, she does not understand the concept of marriage. Iman thinks marriage is like prostitution the relationship between the man and the woman is based on reciprocal interests, "the man pays and the woman serves" (Aboulela, 2019, p. 35).

Therefore, when the three women reach the cottage, and as an act of rebellion, Iman starts to wear customs as a way of acquiring a new identity. She believes that she does not have the chance to choose what she wants, therefore, she thinks that now it is time to choose her path, "It is as if she wants to change her identity because of not being wanted or accepted in the West" (Awajan, 2023, p. 6). Eventually, Iman chooses to take off her hijab, because she wants to break free from traditions, as she thinks that she never chooses to wear it, rather it was forced upon her as an act of habit (Aboulela, 2019, p. 155).

From all that is mentioned above, the three women reach a point where they are overwhelmed, they are struggling to find the right path, and they are stuck in both worlds. However, now they must accept the change, by liberating themselves from the shackles of the past and their old mentalities through transformation. Awajan (2023) states that through their transformation, they can overcome their issues and adopt their new identities as Arab Westerners.

Throughout the novel, Salma, Moni, and Iman offer each other emotional support during their struggles. They listen to one another, provide comfort, and offer advice when needed. Sisterhood challenges the three women's perspectives in how they envision their struggle. Salma, Moni, and Iman, one of them, look at her struggle from her point of view. Thus, they either may succumb to disappointment or take decisions they may regret. However, their bond gives them a platform to confront each other to share mutual awareness. Salma, Moni, and Iman challenge each other by confronting one another about her downfall (Aboulela, 2019).

On the other hand, Salma and Iman confront Moni about the way she handles her life, they blame her for ruining her marriage. Salma says that Moni is prioritizing her son over Murtada and herself. Moreover, Salma advises her to make a balance in order to make her marriage work, while Iman adds that Moni is not a good wife, and pushes her husband away (Aboulela, 2019, pp. 153-157).

Furthermore, Salma and Moni confront Iman about taking her hijab off to fit in Western society. Salma explains to her that taking the hijab off is not a part of tradition, rather it is to protect her from men, and Salma advises her to concentrate on other things to build her new identity, rather than focusing on her beauty (Aboulela, 2019, p. 156). While Moni is suggesting implicitly that it is not by taking off her hijab Iman, it is about keeping their Muslim identity in the West.

Eventually, the heated discussion turns into a quarrel, and every one of them wants to prove her point. Therefore, Salma, Moni, and Iman feel distant from each other every one of them wants her personal space to contemplate her status quo. However, this is the tipping point for Salma, Moni, and Iman, for their transformation. Aboulela draws a vivid picture to the reader of their transformation, the internal feud that they have is represented as a physical metamorphosis. Awajan (2023) and Awajan and Nofal (2025) add that Salma, Moni, and Iman undergo a metamorphosis, transforming into different shapes and creatures. This transformation is a metaphor for the challenges they face as Arabs living in the West and their journey to construct new identities as Arab Westerners. Moni becomes a rolling ball, representing her feeling of being overwhelmed and trapped, Iman transforms into a grotesque mix of mammal and reptile, symbolizing her feeling of being caught between two worlds. Finally, Salma becomes immobilized, losing her strength, which mirrors her feelings of being powerless and trapped by her choices.

At the peak of their transformation, the three women in their unusual state cannot forget each other, so they start calling each other's names, except for Iman who loses her ability to speak. The three of them know that they need each other to overcome this predicament, therefore, when they meet, they "they clung to each other" (Aboulela, 2019, p. 208). Moreover, they feel the power of their togetherness a "unified effort" (p. 208), Salma and Moni say, "We are together" (Aboulela, 2019, p. 208), while Iman draws a large circle around Salma and Moni, and joins them as if she wants to say we are together as well (Aboulela, 2019, p. 208).

Furthermore, Salma's inability to move surrenders her to her friends, whom she trusts to lead her to the right path, and she knows that it is normal behavior sometimes, to ask for help and to be dependent. Salma also stresses the power of sisterhood "Sisterhood was the most valuable and worthy of investments. Moni and Iman might see her as their leader, but she was the one who needed them. She could not now move an inch without them" (Aboulela, 2019, p. 209).

Finally, together they overcome their transformation predicament, and it seems like it is an awakening or a rebirth every one of them now sees the full image, they see now what they are missing, and the moment of epiphany has prevailed. Iman embraces maturity along with accepting her Arab Muslim identity. Aboulela describes Iman's spiritual transformation as "she had grown up. She wore maturity like a cape and it was the best piece of clothing she had ever put on" (Aboulela, 2019, p. 217). Hence, this indicates that Iman now wants to start a new life, to be self-dependent, moreover, to care more about her inner self rather than her beauty. Moreover, Moni starts to love her life, also she changes her perspective, and she wants to handle her life differently, by embracing her son's health condition with love and not as a burden and having more children (Aboulela, 2019).

Finally, Salma stops putting herself down for not passing the medical exam to become a doctor (Aboulela, 2019), she learns to accept her fate that she is a skilled massage therapist, who has a loving husband and beautiful family, therefore she stops chasing the past and embraces the present. In the end, Salma, Moni, and Iman end their journey and return

together holding hands, through their unity they have healed, and they learn to forgive themselves and each other, enabling personal growth and emotional transformation. Aboulela uses the relationship between these women to depict the power of sisterhood and how women can find strength and solace in each other's company, even in unfamiliar and challenging circumstances. Through the characters of Salma, Moni, and Iman, the novel presents a compelling portrayal of sisterhood that resonates with the reader.

IV. CONCLUSION

Both novels portray the complexities of female relationships, particularly the bond between sisters. In *Sula*, the relationship between the protagonist, Sula, and her childhood friend, Nel, evolves and is marked by love, betrayal, and reconciliation. Similarly, *Bird Summons* explores the intricate dynamics between the Salam, Moni, and Iman who embark on a spiritual journey to the Scottish Highlands, delving into their struggles and how they support and challenge each other. Additionally, *Sula* and *Bird Summons* share similarities in their exploration of the complexities of female relationships and sisterhood, portraying the multifaceted nature of women's bonds, the quest for individual identity, and the impact of cultural and societal contexts on these relationships.

ACKNOWLEDGEMENTS

We would like to thank Middle East University in Amman, Jordan, for their financial support granted to cover the publication fees of this research article.

REFERENCES

- [1] Aboulela, L. (2019). *Bird Summons*. London: Hachette UK.
- [2] Abuhasirah, Ramez, & Rasha, Salameh. (2024). Media Framing of Environmental Issues in Jordanian Digital Newspapers. *Middle East Journal of Communication Studies*, 4(1), 1-46.
- [3] Aladaylah, M. (2021). Dislocating Narrativity: Hybridity, Culture and Identity in Leila Aboulela's *The Kindness of Enemies*. *Multicultural Education*, 7(12), 1-6.
- [4] Aladaylah, M. (2023). Together-in-Diaspora: Narratives of Spatial Identities in Leila Aboulela's *Bird Summons*. *Forum for World Literature Studies*, 15(2), 209-221.
- [5] Awajan, N. W., & Al-Shetawi, M. F. (2021). Empowering Muslims in Leila Aboulela's *Minaret* and Mohja Kahf's *The Girl in the Tangerine Scarf*. *International Journal of Arabic-English Studies*, 21(1), 127-144. <https://doi.org/10.33806/ijaes2000.21.1.7>
- [6] Awajan, N. W. (2023). Constructing Arab Western Identities Through Metamorphosis in Leila Aboulela's *Bird Summons*. *World Journal of English Language*, 13(8), 202-209. DOI: 10.5430/wjel.v13n8p202
- [7] Awajan, N., & Nofal, M. (2025). Finding the Third Space in Digital Diaspora in the Narration of Leila Aboulela's *Bird Summons* (2019): A Postcolonial Corpus-Assisted Analysis. *Journal of Arab & Muslim Media Research*. https://doi.org/10.1386/jammr_00103_1
- [8] Bauder, H., & Juffs, L. (2019). 'Solidarity' in the migration and refugee literature: analysis of a concept. *Journal of Ethnic and Migration Studies*, 46(1), 1-20. DOI: 10.1080/1369183X.2019.1627862
- [9] Bolaki, S., & Bröck-Sallah, S. (Eds.). (2015). *Audre Lorde's Transnational Legacies*. University of Massachusetts Press.
- [10] Childs, P., & Williams, P. (1997). *An Introduction to Post-Colonial Theory* (1st ed.). Routledge. <https://doi.org/10.4324/9781315847481>
- [11] Dhavaleswarapu, H. (2016). The Dialectics of Female Friendship in Toni Morrison's *Sula*. *New Academia: An International Journal of English Language, Literature and Literary Theory*, V(II), 95-104. <http://interactionsforum.com/new-academia>
- [12] Fortunati, V. (2019). Chapter One the Heritage of Women Genealogy: Sisterhood and Solidarity Vita Fortunati. In L. M. Crisafulli & G. Golinelli (Eds.), *Women's Voices and Genealogies in Literary Studies in English*. Cambridge University Press.
- [13] Hilton, B. (2011). Frantz Fanon and colonialism: A psychology of oppression. *Journal of Scientific Psychology*, 46-59. http://www.psycencelab.com/uploads/5/4/6/5/54658091/frantz_fanon_and_colonialism.pdf
- [14] Marandi, S. M., & Shabanirad, E. (2015). Edward Said's Orientalism and the Representation of Oriental Women. *International Letters of Social and Humanistic Sciences*, 60, 22-33. DOI: 10.18052/www.scipress.com/ILSHS.60.22
- [15] McLeod, J. (2020). Postcolonialism and feminism. In *Beginning postcolonialism* (second edition) (pp. 197-233). Manchester University Press.
- [16] Mohammad Said Meshaq Omet, A. (2022). *Finding 'home' in Islam: the diversity of Muslim identity in the fiction of Leila Aboulela* [Doctoral dissertation]. Lancaster University. <https://doi.org/10.17635/lancaster/thesis/1624>
- [17] Morrison, T. (1973). *Sula*. New York, NY, Plume.
- [18] Ndayisenga, Z. (2022). Fanon on the Arbitrariness of Using Violence: An Inevitable for both Colonialism and Decolonialization. *Journal of black studies*, 53(5), 464- 484. <https://doi.org/10.1177/00219347221077273>
- [19] Oliveira, N. (2011). *Of Women Bonds: Motherhood, Sisterhood and The Ethics of Care in Toni Morrison's Sula and A Mercy*. Universidade Federal de Minas Gerais.
- [20] Ouakouk, Z., & Bekkari, N. (2023). *What Women Want or Need? Writing the Female Self. Case Study: Spivakian Reading to Toni Morrison's Beloved* [Thesis]. Hamma Lakhdar University of El-Oued.
- [21] Parashar, S. (2016). Feminism and postcolonialism:(En) gendering encounters. *Postcolonial Studies*, 19(4), 371-377.
- [22] Pascual, A. (2020). *Toni Morrison's Sula: A Black Feminist Rereading* [Thesis]. Universidad de Valladolid.
- [23] Sachdev, N. (2020). Feminist Consciousness of Black Women in Afro-American Literature. Anglisticum. *Journal of the Association-Institute for English Language and American Studies*, 9(7), 10-13.
- [24] Said, Edward W. (1978). *Orientalism*. New York, Pantheon Books.

- [25] Sen, T. S. (2021). Sisterhood: A Feminist Motto of Empowerment in Thrity Umrigar's *The Secrets Between Us*. *Lapis Lazuli: An International Literary Journal*, 11(2), 1-10.
- [26] Shovon, A. A. (2022). Friendship and Identity: A Study of Relationship between Nel and Sula in Toni Morrison's *Sula*. *European Journal of English Language and Literature Studies*, 10(8), 62-69. <https://doi.org/10.37745/ejells.2013/vol10n86268>
- [27] Sy, Kadidia. (2008). *Women's Relationships: Female Friendship in Toni Morrison's Sula and Love, Mariama Ba's So Long a Letter and Sefi Atta's Everything Good Will Come* [Dissertation]. Georgia State University. <https://doi.org/10.57709/1059557>
- [28] Tom, A. (2021). Diasporic Sisterhood: Acculturative Stress and the Reorientation of Childhood Female Relationship in Latino Women's Literature. *Navajyoti, International Journal of Multi-Disciplinary Research*, 5(2), 1-12.
- [29] Washington, Steven. (2012). *The African-American Struggle for Equality: Two Divergent Approaches* [Thesis]. Honors College Theses. https://oasis.library.unlv.edu/honors_theses/3
- [30] Yousef, N. T. (2023). Narrative Time in Suzanne Collins' *The Hunger Games*: A Genettian Perspective. *World Journal of English Language*, 13(6). DOI: 10.5430/wjel.v13n6p270
- [31] Yousef, N. T. (2024). Traditional and Postmodern Perspectives: Samiha Khrais's *The Tree Stump* as a Historical Novel. *Dirasat: Human and Social Sciences*, 52(2), 495–508. <https://doi.org/10.35516/hum.v52i2.5165>

Nadera Nader Al-Staity is an MA holder of English language and Literature.

Nasaybah Walid Awajan is an associate professor at the English Language and Translation Department at Middle East University. She earned her PhD from The University of Jordan in English Literature. Awajan has been the Head of English Language and Translation department for two years. She taught as a part time lecturer at The Arab Open University and The University of Jordan. She also worked as an Academic Advisor for English Language teacher students and as the Academic Reading and Writing Lead at Queen Rania Teacher Academy. Her academic work is deeply rooted in the exploration of postcolonialism, Islamophobia, and political theatre, reflecting her commitment to addressing critical social and cultural issues through literary analysis. Awajan's research delves into the intersections of power, identity, and representation, particularly in the context of marginalized voices and the impact of colonial legacies. Awajan has supervised more than 20 MA theses that tackle topics on postcolonial literature, Shakespeare, Contemporary literature and feminist issues.

Sahar Mohammad Tarawneh is an associate professor at the Political Science Department at Middle East University, Amman, Jordan. She earned her PhD from Mu'tah University, Al Karak, Jordan. Her academic work is deeply rooted in Political Sciences and Postcolonialism.

Figurative Language and Aesthetic Value in the Dance Drama *The Search for Sita*

Ni Ketut Dewi Yulianti*

Karawitan Arts Study Program, Faculty of Performing Arts, Institut Seni Indonesia Bali, Denpasar, Bali, Indonesia

Ni Nyoman Kasih

Dance Study Program, Faculty of Performing Arts, Institut Seni Indonesia Bali, Denpasar, Bali, Indonesia

I Wayan Suardana

Crafts Study Program, Faculty of Fine Arts and Design, Institut Seni Indonesia Bali, Denpasar, Bali, Indonesia

Ni Putu Tisna Andayani

Karawitan Arts Study Program, Faculty of Performing Arts, Institut Seni Indonesia Bali, Denpasar, Bali, Indonesia

Sang Ayu Isnu Maharani

English Department, Faculty of Humanities, Udayana University, Denpasar, Bali, Indonesia

Ni Made Ayu Widiastuti

English Department, Faculty of Humanities, Udayana University, Denpasar, Bali, Indonesia

Agus Darma Yoga Pratama

Warmadewa University, Denpasar, Bali, Indonesia

Abstract—*The Search for Sita* is an English-language dance drama adapted from the Ramayana epic, specifically the Aranya Kanda section, which tells the story of *Dewi Sita*'s abduction by Rahwana. This study aims to analyze and describe the figurative language and aesthetic values contained in the dance drama. The approach used is descriptive qualitative, with data obtained through observation of audio-visual recordings of performances and analysis of dance drama scripts, then analyzed using language style theory and aesthetic theory. The results of this study show that the figurative language contained in the dance drama *The Search for Sita* includes metaphors, similes, personification, and hyperbole. Meanwhile, aesthetic values are manifested through unity, complexity, and intensity between the elements of choreography, accompanying music, costumes, and narration. The integration of figurative language and aesthetic values not only strengthens artistic appeal but also presents moral messages and religious values relevant to character education.

Index Terms—*The Search for Sita*, figurative language, aesthetic values, dance drama, *Ramayana*

I. INTRODUCTION

Globalization has expanded the function of performing arts from mere entertainment into a medium for cross-cultural communication and effective cultural diplomacy (Moneta et al., 2025; Mazlan et al., 2025; Suryadmaja & Saearani, 2025). Along with this shift, technological advancement and increased cultural mobility necessitate the revitalization of traditional works in order to maintain their relevance within contemporary socio-cultural dynamics. Such revitalization is commonly realized through innovations in form, language, and modes of presentation that bridge tradition and modernity. A concrete manifestation of this effort is the dance drama *The Search for Sita*, an adaptation of the *Aranya Kanda* episode of the Ramayana epic, performed in English and enriched by the visual aesthetics of Balinese dance costumes.

The linguistic dimension constitutes a crucial element in constructing the dramatic power of a performance. Figurative language, including metaphor, simile, personification, and hyperbole, functions as a stylistic device that intensifies emotional expression and aesthetic resonance (Pranata et al., 2025; Keraf, 2009). In *The Search for Sita*, figurative language does not merely serve as rhetorical ornamentation; rather, it operates as a meaningful vehicle for conveying moral and religious values through poetic imagery. Consequently, examining this linguistic aspect holds substantial academic relevance for understanding the relationship between language use and performative aesthetics within the context of dance drama.

* Corresponding Author. Email: dewiyulianti@isi-dps.ac.id

In addition to its linguistic features, the aesthetic dimension of this work merits scholarly attention due to its synthesis of traditional and modern, local and global, as well as verbal and nonverbal elements within a unified dramatic structure. According to Beardsley (1981), aesthetic value is grounded in three core principles: unity, complexity, and intensity. These principles provide an analytical framework for examining how artistic beauty emerges through the harmonious interaction of elements such as choreography, musical accompaniment, costumes, narration, and stage design. In *The Search for Sita*, these aesthetic components are realized through an integrated visual, musical, and linguistic composition that articulates values of loyalty, sacrifice, and spirituality.

Previous studies offer an important foundation for the present research. Fardeni and Maulina (2025) identify figurative language and imagery as effective means for conveying aesthetic and moral values in literary texts, while Ntelu et al. (2020) demonstrate that figurative expressions deepen emotional meaning and enrich aesthetic layers in Chairil Anwar's poetry. Similarly, Irwanti and Gusthini (2024) reveal that figurative language in *Harry Potter and the Sorcerer's Stone* functions not only as stylistic embellishment but also as a narrative mechanism for constructing imaginative experiences and moral messages. Within the field of aesthetic studies, Mudiasih et al. (2025) show that dance drama contributes significantly to character education through aesthetic experiences that evoke moral awareness. Suryadmaja and Saearani (2025) further emphasize that performing arts in the global era must preserve aesthetic integrity while adapting to cultural transitions and technological transformations. Sumerjana et al. (2025) also highlight that musical aesthetics strengthen emotional engagement and narrative meaning in dance drama performances. Collectively, these studies demonstrate that both figurative language and aesthetic dimensions play a fundamental role in shaping meaning and artistic value.

Despite these contributions, research that explicitly examines the integration of figurative language and aesthetic value within the context of English-language dance drama remains limited. Existing studies tend to focus either on aesthetic analysis or on character education and cultural values, while the interplay between linguistic and artistic dimensions is rarely addressed in a single analytical framework. This gap indicates the need for research that systematically explores the synergy between language and aesthetics in performance.

Therefore, this study aims to analyze and describe the figurative language and aesthetic values embodied in the dance drama *The Search for Sita*. Employing a descriptive qualitative approach, the data are obtained through observation of audio-visual recordings of the performance and analysis of the dance drama script. The data are subsequently examined using theories of figurative language and aesthetic principles. This study contributes to the development of aesthetic and stylistic studies by offering an integrated perspective on how linguistic expression and artistic beauty interact to construct meaning in contemporary dance drama.

II. RESEARCH METHOD

The research method used in this study is a descriptive qualitative approach aimed at analyzing and describing the aesthetic values and figurative language in the dance drama *The Search for Sita*. The research data consists of two types, namely primary data and secondary data. Primary data were obtained through observation of audio-visual recordings of the performance of *The Search for Sita*, while secondary data were collected through analysis of the dance drama script and a review of relevant literature on figurative language and aesthetic values.

Data collection techniques included observation, documentation, and literature study. Observations focused on the dance drama script as well as aesthetic elements of the performance, such as choreography, accompanying music, costumes, and lighting, to identify manifestations of aesthetic values in the staging. Documentation involved the collection of scripts, photographs, and performance notes to support textual and visual analysis. Literature studies were conducted to strengthen the theoretical foundation by referring to Keraf's (2009) theory of figurative language and Beardsley's (1981) theory of aesthetics, along with previous studies related to Ramayana-based performing arts.

The analysis of figurative language was conducted by applying Keraf's (2009) classification, which includes metaphor, simile, personification, and hyperbole. Each dialogue and narrative segment in the script was carefully selected based on its relevance to expressive meaning and symbolic representation. These segments were then coded according to the type of figurative language they represent, allowing for systematic identification and interpretation. The selection of dialogues was based on their frequency of appearance, narrative importance, and representativeness of key themes in *The Search for Sita*, ensuring data relevance and adequacy.

Meanwhile, the analysis of aesthetic values was based on Beardsley's (1981) three criteria: unity, complexity, and intensity. Unity was examined through the coherence between movement, music, narrative, and visual elements; complexity was analyzed by identifying variations and contrasts in choreography, musical patterns, and staging; and intensity was assessed through the expressive strength and emotional impact of the performance. Data analysis was conducted interactively using qualitative procedures that involved data reduction, data display, and conclusion drawing. The findings from observation and script analysis were interpreted descriptively to explain how figurative language and aesthetic values are constructed and conveyed in the dance drama *The Search for Sita*.

III. RESULT AND DISCUSSION

A. Figurative Language Contained in *The Search for Sita* Dance Drama

Figurative language is a form of language use that aims to create aesthetic and emotional effects through deviations from the literal meanings of words (Colston & Gibbs, 2021; Palupi, 2021). According to Keraf (2009), figurative language encompasses several types, such as metaphor, simile, personification, hyperbole, metonymy, synecdoche, euphemism, irony, and allegory. In the dance drama *The Search for Sita*, figurative language is predominantly realized through metaphor, simile, personification, and hyperbole. These four types are discussed in the following section as expressions of aesthetic beauty and semantic intensity, in line with Yulianti and Jayantini (2023), who argue that figurative language not only embellishes discourse but also deepens emotional resonance and spiritual meaning in literary texts.

(a). Metaphor

Metaphors are used to convey meaning through implicit comparisons without using comparative words such as *like* or *as* (Kuzu, 2023; Rahayu et al., 2024). This style of language serves to enrich meaning and present a deeper image by shifting literal meaning to symbolic meaning. Examples of sentences containing metaphors can be found in the following quotes and descriptions.

Beside him, the jewel of Mithila, the radiant Sita. Grace in her steps, kindness in her eyes. Loyal beyond measure, gentle yet strong. She is love made flesh, the light of Rama's path. Sita glances softly toward Rama, her posture calm and serene.

Based on the above excerpt, the metaphors are evident in the phrases “*the jewel of Mithila*” and “*she is love made flesh, the light of Rama's path*.” The expression “*the jewel of Mithila*” does not mean that Sita is a physical jewel, but rather a metaphor that describes her as a symbol of purity, beauty, and the noble values of the kingdom of Mithila. Meanwhile, “*love made flesh*” and “*the light of Rama's path*” are metaphors that personify love and light as the tangible embodiment of Sita, a symbol of love, loyalty, and moral enlightenment for Rama. These metaphors personify love and light as the tangible embodiment of Sita, a symbol of love, loyalty, and moral enlightenment for Rama. This interpretation corresponds with the broader perspective highlighted by Suryati et al. (2023), who note that metaphor functions as a dynamic linguistic device that conveys abstract concepts and emotional values through symbolic representation.

In addition, the use of metaphors can also be found in the following excerpt.

They had no servants, no silk, no gold... yet their hearts were full. Love made every moment rich. Duty made every step sacred.

Based on the above excerpt, the metaphorical style is evident in the sentences “*Love made every moment rich*” and “*Duty made every step sacred*.” Neither of these expressions is meant to be taken literally. Love does not actually produce material wealth, and duty does not literally make a person's steps sacred. Instead, these metaphors illustrate that the true value of life does not lie in worldly luxuries, but in the depth of love and moral awareness in carrying out one's duties. The metaphor “*Love made every moment rich*” describes love as a source of happiness and inner abundance, while “*Duty made every step sacred*” symbolizes devotion to dharma, which makes every action spiritually meaningful. These two metaphors show a transformation from material to spiritual experience, in accordance with the moral teachings in the *Ramayana* story. This interpretation resonates with Rumman and Al Salem (2023), who emphasize that metaphorical expressions often serve as a bridge between the tangible and the transcendent, allowing abstract spiritual values to be conveyed through concrete linguistic imagery, and with Zheng (2021), who notes that metaphorical discourse can effectively communicate cultural and moral ideals across diverse cultural contexts.

(b). Simile

A simile is a figure of speech used to explicitly compare two things by using comparative words such as *as*, *like*, *as if*, or *like* (Keraf, 2009; Salwia et al., 2022). This figure of speech serves to clarify images, emphasize certain impressions, and beautify expressions through similarities between the two objects being compared. Examples of sentences containing similes can be found in the following quotations and descriptions.

From mountain winds and sacred flame, comes Hanuman, son of Vayu, servant of Rama. Joyful. Fearless. Brilliant as the rising sun. His spirit unbreakable. His loyalty is eternal. “Where there is dharma, there is Hanuman.”

Based on the above excerpt, the simile style of language is evident in the phrase “*Brilliant as the rising sun*.” This comparison explicitly uses the word “*as*” to describe *Hanuman's* brilliance, which is likened to the rising sun. This simile not only conveys a visual image of light and radiance, but also evokes symbolic meanings of enlightenment, strength, and spiritual enthusiasm. In this context, *Hanuman* is positioned as a source of moral energy that illuminates the path of truth, just as the sun brings light to the world. The use of simile beautifies the narrative while emphasizing *Hanuman's* character as a heroic and sacred figure. The comparison with the sun confirms his unwavering purity and spiritual strength.

In addition, the use of simile can also be found in the following manuscript excerpt.

It was no ordinary creature. Its fur glowed like firelight. Its eyes held the mystery of forgotten dreams. The forest stood still, watching it.

Based on the above excerpt, the simile style of language is evident in the phrase “*Its fur glowed like firelight.*” This comparison explicitly uses the word like to liken the creature’s fur to the glow of fire. The simile reinforces the visual imagery by presenting a warm, soft, and captivating impression, while also emphasizing the creature’s presence as a figure with extraordinary power that remains in harmony with nature.

In addition, the simile is also reflected in the following excerpt from the text.

Her words, like arrows, struck deep. And though his heart burned with devotion, Lakshmana could no longer stay.

Based on the above excerpt, the simile style of language can be seen in the phrase “*Her words, like arrows, struck deep.*” This sentence explicitly uses the comparative word like to liken Sita’s words to arrows that pierce Lakshmana’s heart. This expression is not meant to be taken literally, but rather describes how Sita’s words have painful emotional power and cause deep emotional wounds. This use of simile not only conveys Sita’s emotional pain vividly but also reflects how figurative comparison can embody cultural and moral values. Such a function aligns with Kardana et al. (2022), who emphasize that simile serves as a linguistic medium for revitalizing local wisdom and strengthening character through emotional and moral expression.

In addition, the use of simile can also be found in the following part of the text.

And then the mask falls. Ravana reveals his true form. No longer the gentle sage, but the fearsome king of demons! His desire burns like fire. His arms reach out, and Sita runs.

Based on the above excerpt, the simile style of language can be seen in the phrase “*His desire burns like fire.*” This comparison explicitly uses the word like to describe Ravana’s burning desire, which is like fire. This simile is not meant to be taken literally, but rather emphasizes the intensity of Ravana’s lust and obsession with Sita. Fire here serves as a symbol of destructive, hot, and uncontrollable power, reflecting Ravana’s basic nature as a figure ruled by worldly desires.

(c). Personification

Personification is a figure of speech that gives human characteristics to inanimate objects, natural elements, or abstract concepts, making them appear alive and capable of behaving like humans (Silaban & Yuhdi, 2023; Imrotin et al., 2022). This figure of speech is often used to strengthen emotional nuances and create a poetic atmosphere in literary works. Examples of the use of personification can be found in the following script excerpts and descriptions.

Alone now. The forest returns to silence. But it has witnessed betrayal. The winds will carry her cry...

Based on the above excerpt, personification is evident in the phrases “*The forest returns to silence*” and “*The winds will carry her cry.*” Both expressions give human characteristics to natural elements, namely the forest and the wind. The forest is described as if it has the awareness to return to silence, while the wind is treated as if it is capable of carrying Sita’s cries, as if it has empathy for her suffering. This personification brings nature to life and makes it feel the tragedy experienced by humans.

The function of personification in this text is to reinforce the emotional and spiritual atmosphere of Sita’s abduction. Nature is not just a passive backdrop, but appears as a silent witness participating in the suffering. The silence of the forest reflects sadness and loss, while the gusts of wind depict the spread of grief that transcends space and time. Thus, personification serves to deepen the aesthetic meaning while presenting a symbolic message that human suffering resonates with the universe, in line with the findings of Yulianti and Jayantini (2023), who assert that figurative language functions as a means of expressing emotions and spirituality in religious texts, and with Pimpuang and Yuttapongtada (2025), who highlight how cultural and spiritual values are reflected through symbolic representations in artistic expressions.

(d). Hyperbole

Hyperbole is a figure of speech that uses exaggerated statements to emphasize meaning and create a dramatic effect (Intang & Jufri, 2025). In literary works and performances, hyperbole is often used to describe strength, greatness, or emotional intensity in an exaggerated manner in order to evoke heroic and monumental impressions. Examples of sentences containing hyperbole can be found in the following script excerpts and descriptions.

From the depths of Lanka, comes Ravana, king of ten heads. A mind unmatched, a will unbending. Pride is his crown. Desire, his sword. He moves with thunder, masked in silk and sin. “Gods may tremble. Kings may kneel. But Ravana... Ravana conquers.”

Based on the above excerpt, hyperbole is clearly evident in the sentence “*Gods may tremble. Kings may kneel. But Ravana conquers.*” This expression exaggerates Ravana’s power by describing how even the gods tremble and kings kneel before him. This statement is not meant to be taken literally, but rather to emphasize Ravana’s extraordinary power and arrogance. This hyperbole serves to highlight the antagonist’s ambitious character and his belief that he is above all other beings. The use of the words trembles and kneel creates a dramatic impression and reinforces the epic atmosphere of the battle between good and evil.

In addition, the use of hyperbole can also be found in the following excerpt from the script.

“I will find you, Sita. Even if I must cross oceans and battle demons, I will bring you home.”

Based on the above excerpt, hyperbole is evident in the statement “*Even if I must cross oceans and battle demons.*” This expression is clearly an exaggerated statement, as crossing oceans and fighting giants are used as symbols of unwavering determination, not as literal descriptions. This hyperbole illustrates the intensity of Rama's love and devotion to Sita, while also emphasizing the moral values of loyalty and courage in the face of any obstacle.

B. Aesthetic Value in *The Search for Sita* Dance Drama

Aesthetics, in the context of performing arts, refers to the study of beauty, artistic perception, and the principles that govern the creation and appreciation of art (Rahimi, 2024; Haseman & Winston, 2010; Hagtvedt, 2020). This includes sensory, emotional, and intellectual responses generated through artistic expression, as well as harmony between form, content, and meaning (Djelantik, 1999; Wicaksandita & Wicaksana, 2025). In this study, Beardsley's (1981) aesthetic framework is operationalized by examining how choreographic structure, musical organization, narration, costumes, and lighting functionally contribute to unity, complexity, and intensity within the performance. An aesthetic analysis of the dance drama *The Search for Sita* shows the integration of traditional and innovative elements, which can be understood through the three aesthetic criteria proposed by Beardsley (1981), namely *unity*, *complexity*, and *intensity*. In terms of unity, as stated by Beardsley (1981), this performance displays internal coherence through the interconnection of elements that support each other's expressive goals. Unity is identified through the practical alignment of movement patterns, musical structure, narrative flow, and visual design that collectively sustain the dramaturgical direction of the performance. The choreography, which is rooted in Balinese dance conventions, moves in floor patterns and transitions (*tandang* and *tangkis*) that are synchronized with the music. This unity is in line with the findings of Fadhlani et al. (2025), who identified four stages of collaborative creativity, namely *exploratory contemplation*, *improvisation*, *expressive collaboration*, and *final touches*, which transform the dancers' bodies into active mediums for cross-disciplinary artistic expression. In addition, *karawitan* music serves as a rhythmic framework, a provider of motifs, and a marker of accentuation (*reyong*, *kendang*, *gong*) that binds the dramaturgical structure.

Balinese puppet-style clothing and makeup preserve visual identity and character typology, while the narrative provides a teleological direction that ensures each segment fulfills its expressive and functional purpose. This coherence generates a sense of closure between scenes, as choreographic motifs, movement phrases, and musical cues continually refer to one another to construct a unified totality of meaning. The resulting unity is not merely the simultaneity of artistic components but an organicity that reflects an internal consistency toward aesthetic and dramatic objectives. As Yulianti and Marhaeni (2021) explain, *unity* in performance is achieved when movement, music, and narrative form a harmonious integration. In this context, the unity of *The Search for Sita* is further reinforced by Beardsley's indicators of functional coherence, the absence of superfluous elements, and the interdependence of meaning throughout the performance. This view is supported by Novianto et al. (2024), who argue that aesthetic unity in Indonesian art often emerges through a dialogue between traditionalism and modernism, maintaining cultural identity amid global aesthetic influences. In a similar manner, Murwanti et al. (2025) emphasize that the principle of unity in artistic practice can also be observed through visual and ecological balance, where each element such as technique, color, and form contributes proportionally to the integrity of the final creation. Therefore, the aesthetic unity of *The Search for Sita* can be understood as a synthesis of traditional continuity, structural harmony, and ecological-cultural awareness that affirms the relevance of performing arts in the discourse of contemporary cultural sustainability. The following figure illustrates one of the scenes from *The Search for Sita*, showing the embodiment of aesthetic balance and cultural symbolism within the performance.



Figure 1. Choreography in the Dance Drama *The Search for Sita*

Beardsley (1981) emphasizes that *complexity* in art is valuable when the diversity of artistic elements produces enrichment without losing coherence or orientation of meaning. In this analysis, complexity is observed through variations in movement motifs, musical textures, character interactions, and symbolic layering that create multiple interpretative levels within the performance. In *The Search for Sita*, complexity emerges through variations of movement motifs developed by meaningful repetition, contrast, and expansion among characters. The differentiation of *Rama*, *Sita*, *Hanuman*, *Ravana*, and *Jatayu* is manifested not only through the quality of movement and energy dynamics but also through spatial patterns and inter-body relations that form multilayered kinesthetic dialogues. This aesthetic complexity reflects a continuous negotiation between tradition and innovation, aligning with the artistic negotiation described by Siswantari et al. (2024), who reveal that dance aesthetics evolve through the harmonization of artistic creativity with cultural and religious values. Similarly, as Karsono (2023) explains, traditional performances and music festivals function as arenas of cultural reproduction and education that sustain collective identity through creative diversity. Seen from this perspective, *The Search for Sita* demonstrates that aesthetic complexity in performing arts extends beyond formal variation to encompass the dialogic interplay of cultural values, learning processes, and intergenerational creativity.

Meanwhile, digital gamelan music contributes to timbral and textural complexity through layers of *ostinato*, *interlocking* patterns, and changes in density that support the emotional dynamics of the scene. This complexity extends to narrative symbolism that maps the conflict between good and evil as an axis of meaning. The results of this analysis are in line with the practice of aesthetic interpretation of symbols in Balinese colossal performances (Kasih & Yulianti, 2025), which shows how the diversity of role iconography deepens meaning. In line with Yulianti et al. (2024), who assert that the aesthetic value of accompanying music is determined by the diversity of musical motifs that give depth of meaning, diversity here operates across elements: musical motifs meet movement motifs, costume typology blends with character markers, and lighting and digital projections thicken the dramatic atmosphere. Thus, complexity does not stop at the accumulation of variations, but is present as an orchestration of differentiation that interprets each other, deepens themes, and affirms the multilayered aesthetic experience of the audience.

Beardsley's (1981) *intensity* framework places expressive quality as a trigger for focused and consistent emotional engagement. Intensity in this study is identified through the management of movement energy, tempo acceleration, spatial compression, musical dynamics, and climactic narrative moments. In the dance drama *The Search for Sita*, intensity is built gradually through the articulation of movement energy, spatial compression, and tempo control so that dramatic peaks such as Sita's abduction and Jatayu's sacrifice present measured emotional outbursts. The intensification of digital *karawitan* music, particularly through *reyong* accents, *gegulet kendang*, harmonization of *calung*, *jegogan*, and *string* melodies, as well as gong resolutions, works as a tension buffer that blends with changes in movement quality and body perspective, producing a strong correspondence between sound and kinesthesia.

Meanwhile, the presence of English narration reinforces communicative intensity by expanding cross-cultural understanding without reducing traditional nuances; this verbal layer clarifies the meaning of references, while the nonverbal layer maintains depth of feeling. In line with Yulianti (2024), who emphasizes the importance of expressive intensity in multicultural choreography, the management of intensity in this work transcends artistic aspects alone and functions as a bridge for intercultural dialogue. Thus, intensity does not appear as a momentary explosion, but as a structured continuity of emotional pressure, ensuring that each dramatic peak has a clear, cohesive, and communicative affective charge for cross-cultural audiences. The following scenes visualize these moments as manifestations of aesthetic energy and narrative power.



Figure 2. Scene of *Ravana* kidnapping *Sita* in the Dance Drama *The Search for Sita*



Figure 3. *Jatayu's Sacrifice* in the Dance Drama *The Search for Sita*

Overall, the dance drama *The Search for Sita* not only presents a combination of figurative language and aesthetics, but also conveys profound moral messages. The moral messages that can be gleaned from *The Search for Sita* are expressed through figurative language that describes spiritual struggle, loyalty, and devotion to *dharma*. Metaphors such as “*the jewel of Mithila*” and “*love made flesh*” emphasize the purity of love and Sita’s steadfastness in upholding loyalty and virtue. Similes and hyperbole reinforce the heroic image and unwavering spirit in upholding the truth, while the personification of nature depicts the involvement of cosmic forces that accompany the spiritual journey of the holy characters. These expressions reflect not only aesthetic beauty but also the wisdom of life that aligns with Vedic perspectives on attaining peace and happiness through self-realization (Yulianti et al., 2023). Overall, the figurative language in this work reinforces the universal message of the power of love, loyalty, and sacrifice as the path to truth and spiritual enlightenment, which is relevant to the formation of a character that is honest, empathetic, and based on the value of truth.

IV. CONCLUSION

The analysis of *The Search for Sita* demonstrates that figurative language and aesthetic principles function in an integrated manner to construct the work’s emotional depth, symbolic meaning, and artistic beauty. The dominant figurative expressions such as metaphor, simile, personification, and hyperbole do not merely embellish the narrative but strengthen the spiritual, moral, and emotional resonance of the story. Metaphors articulate ideals of purity, devotion, and moral enlightenment; similes highlight heroic qualities and emotional intensity; personification animates nature as a witness to human suffering; and hyperbole amplifies the epic scale of conflict and dedication. Together, these linguistic devices support the transmission of universal values such as loyalty, courage, selflessness, and adherence to *dharma*.

In parallel, the aesthetic value of the performance is manifested through the harmonious interplay of choreography, music, costume, narration, and stage design, which collectively fulfill Beardsley’s criteria of unity, complexity, and intensity. Unity is achieved through the structural coherence of movement, music, and narrative; complexity arises from multilayered variations in kinesthetic, musical, and symbolic elements; and intensity is produced through controlled escalation of emotional and dramatic dynamics. These aesthetic dimensions reveal how traditional Balinese performance principles are revitalized through innovative staging, digital musical textures, and cross-cultural narration, resulting in an artwork that is both culturally grounded and globally communicative.

From a broader academic perspective, these findings offer several implications. In linguistics, the study highlights the role of figurative language as a multimodal meaning-making resource that operates beyond textual discourse, extending into performative and embodied expression. In performance studies, the integration of language, movement, music, and visual design demonstrates how aesthetic coherence and emotional intensity are constructed through interdisciplinary collaboration. Furthermore, in the context of cultural adaptation theory, this study shows that the adaptation of the Ramayana into an English-language dance drama can maintain spiritual and moral integrity while enhancing intercultural accessibility, suggesting a viable model for culturally sensitive global performance practices.

Despite its contributions, this study has several limitations. The analysis is limited to a single dance drama production and focuses primarily on textual and performative interpretation without incorporating audience reception or performers’ perspectives. Future research may extend this study through comparative analyses of different Ramayana adaptations, audience reception studies across cultural contexts, or multimodal investigations that combine linguistic, choreographic, and ethnographic approaches. Such studies would further enrich understanding of how language and aesthetics function dynamically within contemporary performing arts.

Overall, *The Search for Sita* illustrates how figurative language and aesthetic construction mutually reinforce one another to produce a performance rich in symbolic meaning, emotional nuance, and cultural significance. Through the synergy of linguistic expression and artistic form, the dance drama succeeds in conveying the timeless moral message

of love, loyalty, sacrifice, and spiritual truth, affirming its relevance as a contemporary work that bridges tradition and modernity while fostering deeper intercultural appreciation.

ACKNOWLEDGEMENTS

The authors convey their deepest appreciation to the Indonesian Institute of the Arts Bali and the Lokananta Dance Community for their dedicated collaboration in the conception and staging of *The Search for Sita*. Profound thanks are also directed to the Swami Vivekananda Cultural Centre (SVCC), the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, as well as to all participating artists and musicians whose creative contributions enriched the production. The authors further express sincere gratitude to the Government of Madhya Pradesh, whose organization of the International Ramayana Mela has not only offered a distinguished platform for artistic expression but also fostered a significant cultural forum where the timeless wisdom of Asian civilizations may be honored and shared with global audiences.

REFERENCES

- [1] Beardsley, M. C. (1981). *Aesthetics, problems in the philosophy of criticism*. Hackett Publishing.
- [2] Colston, H. L., & Gibbs, R. W. (2021). Figurative language communicates directly because it precisely demonstrates what we mean. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 75(2), 228. <https://doi.org/10.1037/cep0000254>
- [3] Djelantik, A. A. M. (1999). *Eстетika: Sebuah pengantar* [Aesthetics: An introduction]. Masyarakat Seni Pertunjukan Indonesia.
- [4] Fadhlani, F., Guntur, G., & Sunarmi, S. (2025). Exploring the creative process of the collaborative artwork "painting in dance": an autoethnographic study. *International Journal of Visual and Performing Arts*, 7(2), 259-270. <https://doi.org/https://doi.org/10.31763/viperarts.v7i2.1782>
- [5] Fardeni, H., & Maulina, I. (2025). Representasi bahasa figuratif dan citraan dalam novel Anjing Mengeong dan Kucing Menggonggong: Kajian stilistika [Figurative language representation and imagery in the novel Anjing Mengeong dan Kucing Menggonggong: A stylistic study]. *JBSI: Jurnal Bahasa Dan Sastra Indonesia*, 5(1), 49–62. <https://doi.org/https://doi.org/10.47709/jbsi.v5i01.5788>
- [6] Hagtvedt, H. (2020). Art and aesthetics. In *Research handbook on luxury branding* (pp. 171-189). Edward Elgar Publishing. <https://doi.org/10.4337/9781786436351.00022>
- [7] Haseman, B., & Winston, J. (2010). 'Why be interested?' Aesthetics, applied theatre and drama education. *Research in Drama Education: The Journal of Applied Theatre and Performance*, 15(4), 465-475. <https://doi.org/10.1080/13569783.2010.512182>
- [8] Imrotin, I., Badrih, M., Famsah, S., & Busri, H. (2022). The Personification Of Figurative Words In Trubus Magazine: A Modern Stylistics Study. *LiNGUA: Jurnal Ilmu Bahasa Dan Sastra*, 17(2), 175-186. <https://doi.org/10.18860/ling.v17i2.17769>
- [9] Intang, A., & Jufri, J. (2025). Metafora dan Hiperbola dalam Novel Laskar Pelangi Karya Andrea Hirata: Kajian Stilistika. *Bahasa: Jurnal Keilmuan Pendidikan Bahasa Dan Sastra Indonesia*, 7(2), 415-425. <https://doi.org/10.26499/bahasa.v7i2.1540>
- [10] Irwanti, R. D., & Gusthini, M. (2024). Analisis Bahasa Figuratif Pada Bab V Novel "Harry Potter And The Sorcerer's Stone". *Jurnal Ilmu Sosial Dan Humaniora*, 2(4), 409-424. <https://doi.org/10.57248/jishum.v2i4.386>
- [11] Kardana, I. N., Rajistha, I. G. N. A., & Satyawati, M. S. (2022). Revitalizing Local Wisdom in Balinese Simile: An Effort to Enhance Balinese Youth Character. *Journal of Language Teaching and Research*, 13(1), 138-144.
- [12] Karsono, K. (2023). Tourism and education intersection: cultural reproduction of Osing community in the Angklung Caruk music festival. *International Journal of Visual and Performing Arts*, 5(2), 106-119. <https://doi.org/10.31763/viperarts.v5i2.1164>
- [13] Kasih, N. N., & Yulianti, N. K. D. (2025). Aesthetic interpretation of waruna symbolism in the colossal dance performance waruna rakta samasta. *International Journal of Visual and Performing Arts*, 7(1). <https://doi.org/10.31763/viperarts.v7i1.1847>
- [14] Keraf, G. (2009). *Diksi dan gaya bahasa (edisi revisi)* [Diction and language style (Revision edition)]. Jakarta: Gramedia Pustaka Utama.
- [15] Kuzu, T. S. (2023). The problem of naming, defining, and teaching of figures of speech: Metaphor, simile, metonymy. *International Journal of Eurasian Education and Culture*, 8(20), 619–630. <https://doi.org/10.35826/ijoecc.674>
- [16] Mazlan, C. A. N., Abdullah, M. H., Hashim, N. S. N., Wahid, N. A., Pisali, A., Uyub, A. I., ... & Hidayatullah, R. (2025). Discovery the intersection of performing arts in cultural tourism: a scoping review. *Discover Sustainability*, 6(1), 14. <https://doi.org/10.1007/s43621-025-00805-w>
- [17] Moneta, A., Hodgson, N., & Fearon, E. (2025). Re-interpreting intangible cultural heritage through immersive live performances to enhance memory-based institutions and foster community engagement. *International Journal of Heritage Studies*, 31(5), 640-661. <https://doi.org/10.1080/13527258.2025.2478612>
- [18] Mudiasih, N. W., Sumerjana, I. K., & Pramanasari, K. D. (2025). Pendidikan karakter dalam drama tari Kesempatan Kedua. *Journal of Innovation Research and Knowledge*, 4(12), 9489–9496.
- [19] Murwanti, A., Hazmi, F., & Kamal, F. (2025). Sustainable batik dyeing with ketapang (terminalia catappa) leaves: a practice-led experimental study on color retention and design innovation. *International Journal of Visual and Performing Arts*, 7(1), 207-217. <https://doi.org/10.31763/viperarts.v7i1.1698>
- [20] Novianto, W., Guntur, G., Faruk, F., & Simatupang, L. (2024). Aesthetic hegemony of the new order: A critical review of cultural politics in Indonesia (1966-1998). *International Journal of Visual and Performing Arts*, 6(2), 162-173. <https://doi.org/10.31763/viperarts.v6i2.1715>
- [21] Ntelu, A., Hintu, E., Yasin, Y., & Supriyadi, S. (2020). Bahasa figuratif dalam puisi-puisi karya Chairil Anwar. *AKSARA: Jurnal Bahasa Dan Sastra*, 21(1), 41–56. <https://doi.org/10.23960/jlg>
- [22] Palupi, M. E. (2021). An analysis of figurative language used in poem of The Echoing Green by William Blake. *JELL (Journal*

- of English Language and Literature) *STIBA-IEC Jakarta*, 6(02), 9-20. <https://doi.org/10.37110/jell.v6i2.131>
- [23] Pimpuang, K., & Yuttapongtada, M. (2025). Language and Culture in Harmony: Buddhist and Chinese Heritage Reflected in Wat Pho's Stone Figurines. *Journal of Language Teaching and Research*, 16(5). <https://doi.org/10.17507/jltr.1605.35>
- [24] Pranata, E. Y., Arifah, I., & Suharto, W. (2025). Makna Simbolis Bahasa Kiasan dalam Antologi Puisi "Selamat Menunaikan Ibadah Puisi" Karya Joko Pinurbo. *Inovasi Sosial: Jurnal Pengabdian Masyarakat*, 2(3), 162-171. <https://doi.org/10.62951/inovasisosial.v2i3.2003>
- [25] Rahayu, S., Purba, J. N., Junaedi, D. P., Fitra, D. N., & Praptiwi, D. (2024). Metafora dalam buku Sehimpun Puisi Ikan-Ikan Terbang, Dari Sungai ke Langit Lengah karya Muhammad Asqalani eNeSTe. *Jurnal Kajian Riset Multidisiplin*, 8(5). Retrieved December 20, 2025, from <https://oaj.jurnalhst.com/index.php/jkrm/article/view/2817>
- [26] Rahimi, M. (2024). The Importance of Aesthetics in Theatre Arts. *Journal of Social Sciences - Kabul University*, 7(3), 159–175. <https://doi.org/10.62810/jss.v7i3.55>
- [27] Rumman, R. N. A., & Al Salem, M. N. (2023). The translation of Quranic metaphorical expressions from Arabic into English. *Journal of Language Teaching and Research*, 14(3), 646-655. <https://doi.org/10.17507/jltr.1403.1>
- [28] Salwia, F., Syahbuddin, S., & Efendi, M. (2022). Analisis Majas Dalam Novel Pasung Jiwa Karya Okky Madasari. *Jurnal Ilmiah Profesi Pendidikan*, 7(4), 2228-2231. <https://doi.org/10.29303/jipp.v7i4.937>
- [29] Silaban, E. M., & Yuhdi, A. (2023). Analisis Gaya Bahasa Personifikasi terhadap Novel "Orang-Orang Biasa" Karya Andrea Hirata. *Lencana: Jurnal Inovasi Ilmu Pendidikan*, 1(3), 43-55. <https://doi.org/10.55606/lencana.v1i3.1771>
- [30] Siswantari, H., Ahimsa-Putra, H., & Maarif, S. (2024). Artistic negotiation, religious and cultural values: An Islamic dance aesthetic in the perspective of Muhammadiyah tarjih decision. *International Journal of Visual and Performing Arts*, 6(2), 234-247. <https://doi.org/10.31763/viperarts.v6i2.1712>
- [31] Sumerjana, K., Mudiasih, N. W., & Pramanasari, N. K. D. (2025). Nilai estetik musik dalam drama tari Kesempatan Kedua. *Prosiding Bali Dwipantara Waskita: Seminar Nasional Republik Seni Nusantara*, 4(1), 325–330. Retrieved December 20, 2025, from <https://eproceeding.isibali.ac.id/index.php/bdw/article/view/556>
- [32] Suryadmaja, G., & Saearani, M. F. T. (2025). Studi Estetika Seni Pertunjukan di Era Global. *Cantata Deo: Jurnal Musik Dan Seni*, 3(1), 71-86. <https://doi.org/10.69748/jmcd.v3i1.339>
- [33] Suryati, N. M., Jimaya, I. K., Kardana, I. N., & Satyawati, M. S. (2023). A Bibliometric Analysis of the Five-Decade Publications on Metaphor. *Journal of Language Teaching and Research*, 14(2), 521-533. <https://doi.org/10.17507/jltr.1402.28>
- [34] Wicaksandita, I. D. K., & Wicaksana, I. D. K. (2025). Śiva Naṭarāja: Konsep Estetika Hindu dalam Seni Pertunjukan Wayang. *Panggung*, 35(3), 419-441. <https://doi.org/10.26742/panggung.v35i3.2693>
- [35] Yulianti, N. K. D., & Jayantini, I. G. A. S. R. (2023). Textual Metafunction of Cynicism in Srimad Bhagavatam and Its Translation into Indonesian. *Journal of Language Teaching and Research*, 14(5), 1244-1254. <https://doi.org/10.17507/jltr.1405.12>
- [36] Yulianti, N. K. D., & Marhaeni, N. K. S. (2021). Analisis Nilai Estetika Pertunjukan Wayang Kulit Cenk Blonk Dalam Lakon "Tidak Cukup Hanya Cinta". *Panggung*, 31(2). <https://doi.org/10.26742/panggung.v31i2.1593>
- [37] Yulianti, N. K. D., Ardini, N. W., & Darmayuda, I. K. (2023). Wisdom to attain happiness in Vedic culture perspective. *Linguistics and Culture Review*, 7(1), 1-9. <https://doi.org/10.21744/lingcure.v7n1.2234>
- [38] Yulianti, N. K. D., Sariada, I. K., & Marajaya, I. M. (2024). The Aesthetic Value of the Accompaniment Music of the Dance Drama 'The Blessing of Siva-Visvapujita' | Nilai Estetika Musik Iringan Drama Tari "The Blessing of Siva-Visvapujita". *GHURNITA: Jurnal Seni Karawitan*, 4(3), 265-272. <https://doi.org/10.59997/jurnalsenikarawitan.v4i3.3724>
- [39] Yulianti, N. K. D. (2024). *The Blessings of Siva-Visvapujita: Kisah Pohon Suci Tulasi dalam Koreografi Multikultur*. PT Nilacakra Publishing House.
- [40] Zheng, Y. (2021). The Effectiveness of Xi Jinping's Metaphorical Discourse in Cross-Cultural Communication. *Journal of Language Teaching and Research*, 12(6), 1051-1055. <https://doi.org/10.17507/jltr.1206.21>



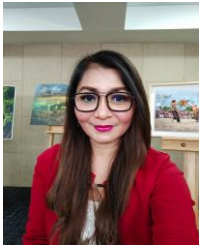
Ni Ketut Dewi Yulianti, born in Karangasem, Bali, on July 23, 1973, earned her Doctorate in Linguistics from Udayana University, Denpasar, Indonesia, in 2016, and was inaugurated as a Professor in Cultural Linguistics in 2024. She has received several research and community service grants from the Ministry of Education, Culture, Research, and Technology. She currently teaches English and is also active as an instructor of Bahasa Indonesia for Foreign Speakers (BIPA). She frequently serves as a speaker and presenter at local, national, and international academic conferences. She is a member of academic and cultural communities engaged in language, arts, and cross-cultural studies. Her current research interests include linguistics, cultural linguistics, language education, and performing arts studies.



Ni Nyoman Kasih, born in Mas, Ubud, Gianyar, Bali, on October 20, 1964, has been a lecturer at the Indonesian Institute of the Arts (ISI) Bali since 1993. She completed her Master's degree in Art Creation at ISI Yogyakarta in 2002. She teaches in the Dance Study Program and has also taught across faculties within ISI Bali. Her research includes *Lukat: Transformation of Reality into Digital Art* and *Daksa Curses Lord Siva*. In addition to research, she has written several academic papers such as *Strategi Pembelajaran Gerak Dasar Tari Bali untuk Anak-Anak di Sanggar Tari Warini*, *Bentuk dan Fungsi Tari Telek Melampahan di Nusa Lembongan*, *The Sang Hyang Kelor Ritual in Lembongan Village, Klungkung, Bali*, and *Aesthetic Interpretation of Waruna Symbolism in the Colossal Dance Performance Waruna Rakta Samasta*.



I Wayan Suardana, born in Petulu, Bali, on December 31, 1963, is an Associate Professor at the Indonesian Institute of the Arts (ISI) Bali specializing in Creative Art Craft. He earned his Doctorate in Art Creation and Study from ISI Yogyakarta in 2017. He has received several recognitions, including the Certificate of Competence in Art Curation (2024), the Capacity Building Program in Creative Economy from Indonesian AID (2024), and an award as keynote speaker in the Plastic Waste Processing Seminar (2025). He actively participates in national and international art exhibitions such as the IVCE 4 International Visual Culture Exhibition (2022), Bali Mega Rupa (2022–2025), B_Game Bali Global Art Map (2024), and the International Art and Craft Exhibition *Jejak Rasa* in Bandung (2025). His works focus on contemporary craft rooted in tradition, emphasizing aesthetic, material, and cultural narrative exploration.



Ni Putu Tisna Andayani, born in Denpasar, Bali, on May 29, 1978, has been a lecturer at the Karawitan Study Program, Faculty of Performing Arts, Indonesian Institute of the Arts (ISI) Bali since 2005. She earned her Master's degree in Applied Linguistics with a concentration in Translation from Yogyakarta State University in 2014. In addition to teaching and researching in linguistics and translation, she is active as an instructor of Bahasa Indonesia for Foreign Speakers (BIPA) and often serves as a moderator, interpreter, and master of ceremony at academic and cultural events.



Sang Ayu Isnu Maharani, a lecturer at the English Department, Faculty of Humanities, Udayana University, is a certified Public Relations assessor and English Medium Instruction (EMI) instructor. She earned her Master's degree in Applied Linguistics from Udayana University and has been actively teaching language skills, literature, and linguistics. Her research interests include sociolinguistics, anthropological linguistics, and translation studies. She has served as a public speaking coach for students and founded the *Master of Ceremony (MC) Community* of the Faculty of Humanities in 2018. She has assisted numerous journal writings as a proofreader and translator, including the *Master Plan of Gianyar (2015–2019)* and *Gianyar as the Heritage City*. From 2011 to 2017 she was the Coordinator for International Students at the International Office of Udayana University, and in 2019 she was appointed as Head of the Language Center. She remains active in community service and academic collaborations at local and international levels.



Ni Made Ayu Widiastuti is a lecturer at the English Department, Faculty of Humanities, Udayana University, Bali, Indonesia. She holds a doctoral (2024) degree in Linguistics from Udayana University. Her current research interests include Corpus Linguistics, and English-Indonesian-English Translation Studies. She has published several articles on these topics in international journals.



Agus Darma Yoga Pratama is one of the lecturers at the Faculty of Letters, Warmadewa University, who completed his Doctoral Degree (S3) at Udayana University in 2016. Translation is one of the focuses of research that has been published in several seminars and journals, in addition to other linguistic fields such as language teaching, discourse studies, and applied linguistics with a focus on the interrelation between language, culture, and communication.

Language and Digital Content in Saudi E-Commerce: An Analytical Review

Mohammad Abdulaal*

Department of English, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Kingdom of Saudi Arabia

Mohammad Aldawoody

Department of English, Faculty of Arts, Port Said University, Port Said, Egypt

Abstract—The gradual rise of e-commerce in Saudi Arabia has generated new needs for linguistically and culturally qualified digital interaction. This analytical review investigates how the use of language and the design of digital data overlap with competent interactional conventions in electronic trade environments in Saudi Arabia. Employing an integrative methodology that synthesizes both peer-reviewed research studies and industry analyses, the review explores how Arabic and English are utilized across e-commerce platforms. The analytical review also focuses on the implications of localization, quality of translation, and intercultural communication for professional practices and training. The results suggest that culturally adjustable language tactics promote consumer confidence. However, deficiencies are noticed in the preparation of content for bilingual and culturally perceptive communication. The review concludes that specialized cross-cultural training and localized language strategies are important to enhance clarity, brand reliability, and user experience in Saudi e-commerce. Future research should further investigate the correlation between linguistic adaptation and digital consumer conduct to guide professional communication training in multilingual trade sectors.

Index Terms—digital discourse, intercultural pragmatics, localization, bilingual use, cultural adaptation

I. INTRODUCTION

E-commerce in Saudi Arabia has swiftly developed as a substantial catalyst of the domestic economy. This rising growth is in accordance with Vision 2030's electronic transformation targets. It not only reflects technological advancement but also presents significant linguistic and cultural challenges. It is due to the fact that commercial platforms must balance Arabic, the official native language, with English, the primary medium for international trade. In this context, language serves not only as a device for information delivery but also as a construct of trust, identity, and consumer engagement. It makes the design and localization of digital content a key factor for online success and continuity (Addi & Teimi, 2025; Aguirre & Domahidi, 2021; Boginskaya, 2022).

From a sociolinguistic lens, effective language use in digital trade entails more than accurate translation. It entails cultural adjustment and sensitivity to users' anticipations. Studies indicate that many international and local e-commerce platforms face limitations in Arabic localization. For instance, Alharbi et al. (2023) pointed out that while global trademarks like Apple and Huawei uphold high global standards, their Arabic interfaces often showcase partial localization and insufficient socio-cultural alignment. Similarly, Omar et al. (2022a) identified translation inconsistencies, weak cultural adaptation, and accessibility issues as recurring challenges that can shake consumer trust and engagement.

Language choice additionally impacts trustworthiness and consumer behavior. Ansari (2021) points out that Arabic translation in electronic shopping questionnaires is substantial for reflecting local opinions of trust and interface effectiveness. Earlier research studies emphasize that Saudi electronic trade platforms often underperform global competitors regarding content quality and customer orientation (Alotaibi, 2013). These linguistic and content-related factors can determine whether consumers perceive online retailers as professional, transparent, and culturally aligned (Anspoka & Eglite, 2019; Casañ Pitarch, 2017; Wood, 2007; Kriukova, 2022).

Professional communication in digital commerce thus requires highly specialized competencies. Linguists, digital advertisers, and web developers should merge linguistic precision, cultural localization, bidirectional interface design, and audience demands. Despite rapid advancement in information and communication technology infrastructure and e-commerce adoption (AlGhamdi et al., 2012), there is limited or little research on the development of professional communication techniques or training schemes within Saudi e-commerce institutions. Chicherina and Chicherina (2021), Chung (2012), and Omar et al. (2022b) underline the need for systematic evaluation of localization quality and for enhancing bilingual expertise among digital content collaborators.

Additionally, Saudi shoppers always expect bilingual or Arabic-prioritized encounters that embody national identity and cultural conventions. Research on social trade indicates that language choice significantly molds users' reliability,

* Corresponding Author.

sense of community, and engagement with digital brands (Alotaibi & Aljaafari, 2024; Herzallah, 2022; Contentech, 2025). However, the linguistic aspect of e-commerce remains understudied compared with technical and logistical aspects.

Therefore, this analytical review aims to integrate current research on language practice and electronic content in Saudi electronic trade. It also aims at defining the major tendencies, barriers, and implications for professional communication and training. To attain this aim, this analytical review endeavors to answer the following questions:

- [1] Which language do Saudi and international electronic trade portals use to communicate with shoppers?
- [2] What are the linguistic and cultural barriers that encounter the localization of Saudi electronic commerce data?
- [3] How can the language of content layout impact client confidence and buying choices in the Saudi electronic trade portals?

By incorporating linguistic, cultural, and professional communication dimensions, this review puts forth insights relevant for scholars, practitioners, and policymakers.

II. LITERATURE REVIEW

A. Language as a Component of Digital Communication

In online commerce, language functions as an essential channel of professional interaction that reflects both credibility and trust. Theories of digital discourse point out that electronic texts do not simply deliver information but form consumers' perceptions via linguistic tone, lexical choice, and multimodal elements (Kress, 2010). In the Saudi Arabian setting, language has an emblematic role, where Arabic acts as an indicator of national identity while English represents global inclusion. According to Crystal (2011), English serves as an international lingua franca, facilitating e-commerce and technology. But Saudi consumers usually interpret it through culturally specific structures (Koşar, 2023; Korenyugina, 2025; Lee, 2014; Woods, 2018; Al-Khateeb, 2023; Al-Wossabi, 2016).

Research studies on Arabic digital communication (Al-Said & Al-Khalifa, 2019) show that the sociolinguistic intricacy of Arabic posits obstacles for localization and translation. The complexity of the Arabic language is due to its diglossic nature. This corresponds with Halliday's (2004) systemic functional model. Therefore, the efficiency of Saudi online commerce communication relies not only on linguistic precision but also on the correspondence of textual options with cultural expectations and digital literacy standards (Li, 2025; Lilac, 2009; Witherspoon, 1980).

B. Localization and Cultural Adaptation

The notion of localization exceeds the concept of translation to encompass the sociocultural, technical, and aesthetic adaptation of online content to designated users. Esselink (2000) describes localization as a process that guarantees that a product "looks and feels as though it were developed particularly for the target market." In Saudi e-commerce, localization involves converting not only the language of websites and applications into Arabic but also adjusting currency, imagery, design direction (right-to-left interface), and even communication styles to match cultural norms.

Omar et al. (2022a), Chernyavskaya and Zharkynbekova (2024), and Ibrahim et al. (2013) concluded that Arabic electronic commerce websites show limited socio-cultural adaptation. It caused inconsistent lexical items and operability problems. Like Omar et al. (2022b), Singh and Pereira (2005) pointed out that culturally tailored websites generate higher user reliability. Hofstede's (2001) socio-cultural dimensions model also offers an interpretive viewpoint: Saudi Arabia ranks high on communalism and ambiguity avoidance. It implies that localized digital content must stress trust, community belonging, and transparent guidance to decrease assumed risk (Masango, 2006; Müller & Freudenthaler, 2022).

Hence, social and cultural adaptation signifies an integration of linguistic accuracy and cross-cultural proficiency. According to Alharbi et al. (2023), even international enterprises, such as Apple and Huawei, face challenges to adapt their Arabic digital platforms to Saudi cultural norms. The gap between international brand identity and local interaction suggests that localization needs both social and cultural knowledge. This aspect is usually addressed in translation studies and intercultural communication training.

C. Intercultural Communication and Consumer Trust

The intercultural communication theoretical model offers a sound basis for comprehending how linguistic and cultural insights impact credibility and engagement in electronic commerce. Hall's (1976) differentiation between dependent- and independent-context cultures illuminates Saudi consumers' preference for implicit communication styles, where context and shared cultural understanding carry significant meaning. In context-based cultures like Saudi Arabia, visual signs, politeness strategies, and culturally meaningful language play substantial roles in reflecting respect and reliability (Suvorova et al., 2024; Sun, 2017; Mouhoubi-Messadh & Khaldi, 2022).

Consumers' confidence in digital settings has been widely examined (Gefen, 2000; Kim et al., 2008; Rahman, 2022). The findings suggest that linguistic and cultural correspondence promotes users' engagement in online transactions. In Saudi Arabia, Ansari (2021) observed that bilingual communication—combining Arabic and English—can promote both accessibility and credibility, especially among younger, digitally fluent consumers. This dual-language strategy supports linguistic inclusivity while maintaining international market appeal.

D. Professional Communication and Training Implications

From an occupational interactional point of view, influential online trade requires competencies that combine linguistic precision, cultural consciousness, and digital fluency. According to Bargiela-Chiappini (2014), theories of professional discourse stress that professional communication takes place in multilingual and multimodal digital settings. For Saudi practitioners (i.e., translators, marketers, and web designers), this suggests the need for training that connects applied linguistics, intercultural discourse, and media studies. Current literature points to a skills gap in localization and intercultural digital content production across the Arab world.

According to Alotaibi and Aljaafari (2024), the widespread of Saudi e-commerce reflects strong technological adoption. But it also shows limited readiness among translators and marketers to generate culturally adapted Arabic content. By the same token, Alharbi et al. (2023) and von Bitter and Turley (2016) highlight the absence of standard frameworks for assessing language quality and cultural fit in Saudi e-commerce websites. Training programs focusing on digital language competence are thus essential for advancing the country's digital economy goals. To address this, Saudi higher education curricula should move beyond traditional translation instruction to include digital content production, multimodal communication, and localization technologies. Concurrently, teacher education programs must equip instructors with interdisciplinary skills to effectively train students in applying linguistic, cultural, and digital competencies in professional e-commerce contexts.

E. Theoretical Integration

The theoretical foundation for this review article is built upon three interconnected perspectives. Together they allow for a multidimensional investigation of the Saudi electronic commerce environment. Systemic Functional Linguistics (SFL) is highly grounded on Halliday (2004). The SFL describes language not as a set of static rules but as social action fundamentally shaped by its context and purpose. It facilitates the analysis of linguistic choices in online interactions. The SFL is combined with Intercultural Communication Theory (Hall, 1976; Hofstede, 2001). This theory offers the framework for explaining how inherent cultural tendencies in the Arab context affect the interpretation of meaning. Also, it explains users' responses and engagement with digital content. Finally, Localization and Digital Discourse Theory was developed by Esselink (2000) and Kress (2010). It explains the shifting interaction between linguistic choices, visual design aspects, and localization methods in digital environments. By integrating these theories, the review is prepared to assess Saudi e-commerce as a virtual space where linguistic choices, cultural norms, and professional practices intersect. The three theories together offer a solid basis for assessing how language actively shapes consumer experience and social trust (Xia, 2024; Yousef et al., 2012; Werner, 2022).

III. METHODOLOGY

This article embraces an analytical review approach to explore how language use and digital content practices overlap within the Saudi electronic trade. Instead of compiling primary data, this review methodologically integrates the current scholarly literature, official reports, and commercial sector analyses to identify linguistic and cultural aspects affecting digital communication in online retail contexts. The aim is to synthesize scattered findings, assess emerging trademarks, and underscore implications for professional communication and intercultural training.

A. Research Design

The study implements an explanatory analytical review design, appropriate for examining interdisciplinary themes that integrate linguistics, communication, and online trade. This approach adopts a qualitative interpretation of research findings across the relevant fields, stressing conceptual correlations rather than statistical extrapolations. The review endeavors to portray how language selection, localization procedures, and cultural accommodation mold consumers' perceptions, user engagement, and professional communication standards in Saudi Arabia's growing online commerce.

B. Data Sources and Materials

Data were collected from conference proceedings and scholarly papers published in peer-reviewed journals between 2012 and 2024. This timeframe was chosen to align with Saudi Arabia's rapid digital transformation under Vision 2030. Investigations were performed in major academic repositories comprising Scopus, Web of Science, and ScienceDirect to retrieve pertinent articles. Official and commercial reports from sources such as the Saudi Ministry of Communications and Information Technology were also reviewed to retrieve more contextual data. The search methodology integrated key terms and Boolean operators, such as ("Saudi Arabia" AND "e-commerce") AND ("language use" OR "digital content" OR "localization" OR "translation quality" OR "cultural adaptation" OR "bilingual communication").

This search yielded 110 sources, which were reviewed according to eligibility and exclusion standards. The criteria for study selection ensured a focus on the intersection of language, communication, and digital commerce within the Arab world. For a study to be selected, it needed to specifically explore language use, interaction practices, or cultural customization within the Saudi or pan-Arab digital markets. Additionally, valid studies had to explore linguistic practices, multilingualism, or cross-cultural communication. To ensure academic standards, only peer-reviewed sources from recognized institutions were considered, and they had to be published in English and accessible in full text.

Alternatively, studies were disregarded if their main emphasis was on areas ancillary to the central theme, such as solely technical, financial, or logistic aspects of e-commerce that lacked linguistic or cultural analysis. Non-scholarly

materials, including editorials and general commentaries, were rejected, as was data that duplicated information found in other retained sources. Studies were also disregarded if they investigated non-Arab settings with no direct relevance to the Arabic-language markets. After applying these detailed inclusion and exclusion parameters, the final corpus for analysis included 32 key research studies and 10 supplementary reports.

C. Instruments and Participants

This study did not involve human participants, as it adopted an explanatory analytical review design rather than empirical data collection. The participants of the study are represented by previously published scholarly articles and institutional reports that address language use, localization, and cultural adaptation in Saudi digital commerce. These sources constituted the units of analysis. The primary research instrument was a structured document analysis protocol. This instrument guided the systematic identification, screening, and evaluation of relevant literature. Peer-reviewed journal articles, conference proceedings, and official reports published between 2012 and 2024 were included. The materials were retrieved from recognized academic databases and authoritative national sources. Clear inclusion and exclusion criteria ensured thematic relevance and methodological rigor. The selected texts were qualitatively analyzed to examine linguistic, cultural, and professional communication practices in online retail contexts.

D. Analytical Procedures

The included materials underwent thorough qualitative thematic analysis. This included reviewing each selected research study for its specific goals, methodological approach, and key findings as they related to three key domains of investigation. The first domain centered on language and localization, investigating practical issues such as bilingual interfaces, translation fidelity, and terminology consistency across platforms. The second domain, Cultural Adaptation and User Perception, concentrated on how design choices impacted consumer credibility, engagement levels, and the overall cultural sensitivity of digital settings. Finally, the third domain investigated professional communication and training implications, particularly evaluating the competence and required skills of translators, and digital content specialists.

The following analytical stage included synthesizing themes that surfaced across the corpus of research. The findings were compared to detect areas of commonalities, discrepancies, and unexplored gaps. Thematic categorization focuses on repeated major challenges. These obstacles included the widespread problem of insufficient Arabic localization and the use of varying translation criteria. These established patterns were critically assessed and demonstrated employing representative research studies from the corpus, such as Alharbi et al. (2023), Omar et al. (2022a), Alotaibi and Aljaafari (2024), Frank et al. (2001), and Grace et al. (2019).

E. Reliability, Validity, and Ethical Considerations

To ensure analytical standards, the review adopted a repetitive reading and cross-checking process. Each selected study was read two times. The first time was for a general understanding of the study. The second time of reading was for in-depth analysis and grouping of sociolinguistic aspects. Findings were corroborated with industry data and government reports to contextualize scholarly perspectives. Reference management applications (Zotero) were utilized to organize references and prevent citation mistakes. Because this review depended solely on released materials available in the public sphere, no ethical clearance was needed. However, all intellectual rights were credited through accurate referencing and compliance with APA 7th edition referencing standards. No personal or confidential data were employed.

IV. RESULTS

A. Overview of the Reviewed Studies

This analytical review includes refereed articles, reports, and governance texts issued between 2012 and 2024. They focus on language use, localization quality, and cultural interaction in Saudi Arabia's electronic commerce. This corpus includes research articles and commercial analyses, focusing on customers' linguistic experiences and choices.

A consolidation of these sources includes three repeated research areas: the bilingual duality of digital retail platforms, the socio-cultural adaptation and localization problems that form users' experiences and content suitability, and the continuing shortcomings in professional communication practices and training needed for effective multilingual digital involvement.

The reviewed research studies show a set of interconnected tendencies that form how language, culture, and occupational practices intersect in the Saudi electronic commercial settings. These tendencies underscore the linguistic hybridity of online interaction. Also, they highlight the degrees to which digital websites succeed or fail in adapting their content to local cultural norms. Additionally, the corpus underscores the centrality of consumer trust, demonstrating that linguistic accuracy and cultural resonance significantly impact users' perceptions. At the same time, ongoing gaps in professional training and localization expertise show structural challenges that continue to influence the quality of Saudi digital commerce. The table below combines these thematic insights and situates them within key representative studies.

TABLE 1
MAJOR RESEARCH THEMES AND REPRESENTATIVE STUDIES

Theme	Key Findings	Representative Studies
Bilingual Communication (code-switching)	Saudi e-commerce platforms display code-switching between Arabic and English, reflecting both national identity and international modernity.	Al-Said & Al-Khalifa (2019); Crystal (2011)
Localization and Cultural Resonance	Localization caliber varies; some international brands show partial tailoring, while local websites achieve higher cultural alignment.	Alharbi et al. (2023); Omar et al. (2022b)
Consumer Trust and Linguistic Accuracy	Language precision, politeness conventions, and cultural awareness strongly impact consumer trust.	Ansari (2021); Gefen (2000); Kim et al. (2008)
Professional Training and Skill Gaps	Lack of localization and cross-cultural training restricts online retail competitiveness.	Alotaibi & Aljaafari (2024); Esselink (2000); Bargiela-Chiappini (2014)

B. Language Use and Bilingual Practices

Saudi online retail platforms operate in a dual-language environment. Some studies, such as Al-Said and Al-Khalifa (2019) and Crystal (2011), show that the Arabic language is employed for relational interaction and socio-cultural alignment. English, on the other hand, is used for functional and technical purposes. This linguistic duality implies both cultural pride and global orientation. It is in alignment with Vision 2030's modernity discourse.

However, dual-language practice usually causes irregularities in language use and translation caliber. Omar et al. (2022a) concluded that only 42% of Arabic-language online retail websites met acceptable coherence and fluency standards. Errors in terminology, sentence structure, and register reduce clarity and undermine user trust.

Furthermore, some websites display hybrid scripts, combining Arabic titles with English product descriptions or customer reviews. While this mixture broadens accessibility, it can create cognitive strain for monolingual Arabic users. This finding corresponds to international online retail studies showing that linguistic mismatch between user anticipations and site language decreases purchase intention (Singh & Pereira, 2005).

C. Localization and Cultural Adaptation

Localization is substantial to electronic interaction caliber. Alharbi et al. (2023) drew a comparison between Arabic interfaces of Apple and those of Huawei. They try to spotlight the differences in terms of cultural adaptation. On the one hand, Huawei's Arabic website integrated culturally specific images and a restrained layout. Apple, on the other hand, retained an international neutral form. Saudi users viewed Huawei's localized website as more reliable.

By the same token, Alotaibi and Aljaafari (2024) noted that online commercial startups employ highly polite Arabic forms, greetings, and culturally relevant images. These findings are in alignment with Hofstede's (2001) model of high-context interaction. According to this model, semantic content relies entirely on common cultural perception rather than direct verbal information. Despite improvements, localization remains uneven across international and domestic platforms. Omar et al. (2022b) and Shi and Tian (2019) related this to the lack of trained localization professionals who combine linguistic competence with technical expertise in website designing. This professional gap restricts the consistent application of localization standards across Saudi digital content.

TABLE 2
LOCALIZATION AND COMMUNICATION CHALLENGES IN SAUDI E-COMMERCE

Category	Observed Issues	Impact on Communication	Supporting Studies
Translation Caliber	machine-translated Arabic content; inconsistent terminologies.	Reduces clarity and credibility.	Omar et al. (2022a); Singh & Pereira (2005)
Cultural Sensitivity	Lack of culturally suitable visuals.	Low or no user engagement; perceived cultural discrepancy.	Alharbi et al. (2023); Hofstede (2001)
Bilingual Imbalance	English dominance and mixed scripts.	Confusion among monolingual customers; fragmented identity.	Al-Said & Al-Khalifa (2019)
Professional Competence	Limited or no training in localization and cross-cultural communication.	Unclear branding and poor content resonance.	Bargiela-Chiappini (2014); Alotaibi & Aljaafari (2024)

D. Intercultural Communication and Consumer Trust

Cross-cultural communication models throw much light on how language and layout form users' credibility. Hall's (1976) high-context culture framework clarifies why Saudi consumers prefer indirect signals such as politeness and tone. Ansari's (2021) survey of 400 Saudi consumers discovered that 72% preferred dual-language interfaces. However, cultural familiarity most strongly predicted trust and repeated purchases from the same place.

Gefen (2000) and Kim et al. (2008) reinforce this by connecting perceived similarity and linguistic congruence with online trust formation. For Saudi users, commercial interfaces reflecting Islamic norms with a respectful tone and conservative images mark trustworthiness. In contrast, overly Westernized elements may dissatisfy or confuse local audiences. The emotional resonance of culturally adapted language thus becomes a key determinant in Saudi digital branding.

E. Professional Training and Competence Gaps

A major theme arising throughout research studies is the lack of practitioners trained in dual-language digital discourse. As online retail grows, the demand for professionals in localization, translation, and cross-cultural marketing increases. However, Alotaibi and Aljaafari (2024) point out that Saudi higher education still stresses the traditional translation rather than electronic content production. This discrepancy leads to a lack of systematic localization norms. Further, it impedes market competitiveness.

Bargiela-Chiappini (2014) sheds light on the need to incorporate intercultural discourse into business education to cultivate linguistic and cultural awareness in professional contexts. Calls for reform include creating certification programs for localization specialists and promoting collaboration between linguistics departments and industry stakeholders (Esselink, 2000). Such initiatives would strengthen Saudi Arabia's digital workforce and align language training with Vision 2030's economic diversification goals.

Contemporary literature emphasizes diverse, dynamic, and emerging trends forming the Saudi online markets, mainly motivated by an emphasis on enhanced cultural resonance and technological incorporation. There is a rising need for an Arabic digital identity. Such an identity makes Arabic interfaces the first preference for customers. This focus is complemented by the need for dual-language inclusivity. Such bilingual accessibility ensures a smooth incorporation of English content to serve the non-national community.

However, this bilingual inclusiveness should not undermine the Arabic cultural identity of the platform. The role of digital tools is notable. It can be noticed in the artificial intelligence-based localization trend that is often supported by human post-editing to guarantee both accuracy and appropriate cultural tone (Omar et al., 2022b). Furthermore, there is an increase in academic-industry collaboration, where universities are partnering with companies to embed practical localization modules into applied linguistics and related curricula (Alotaibi & Aljaafari, 2024). Overall, interaction strategies are shifting towards user-based layouts, transitioning from merely product-focused messaging.

V. DISCUSSION

The analytical integration of the reviewed literature results in three main, interrelated results. As a matter of fact, these results reflect the future linguistic requirements of the Saudi digital markets. The first result shows that there is a significant correlation between the quality of linguistic and cultural localization and consumer metrics. It indicates that proper localization impacts user trust and brand reliability. The second result reveals the lack of trained practitioners who have linguistic proficiency, cultural comprehension, and technological mastery. All these skills are needed for influential localization and proper digital discourse. Finally, the third result sheds light on the growing digital space as developing a special bilingual identity. Such a space incorporates Arabic cultural signals and values with English-mediated competency in digital discourse.

The results of this analytical review reveal that the linguistic and cultural dimensions of digital discourse are significant to the success of digital retail. Language, as the vehicle of meaning and self-representation, plays a twofold role: it serves as an operational medium for information transfer and as a metaphorical resource for cultural identity and trust-building. The presence of Arabic and English in Saudi online markets reflects the kingdom's local self-image and its linguistic integration into the digital economy.

As Alharbi et al. (2023) and Alotaibi and Aljaafari (2024) pointed out, impactful localization techniques do more than translate words. They align messages with cultural expectations, social norms, and visual standards. This alignment is significant in high-context cultures like Saudi Arabia. It is due to the fact that meaning in high-context cultures is usually expressed in an indirect way via tone, metaphors, or shared beliefs (Hall, 1976; Peniro & Cyntas, 2019; Rahman, 2014). When online retail content falls short of conveying these subtleties through overly literal translation, it discourages users who view such options as indicators of cultural insensitivity.

The dominance of bilingual retail websites reflects an emerging linguistic identity in digital discourse. This bilingualism represents a mixed communicative model rather than a clash between Arabic and English. Such a model incorporates international availability and cultural integrity. This phenomenon enhances Crystal's (2011) notion of "digital bilingualism." This notion states that users fluidly move through multiple linguistic systems, relying on context and purpose. In Saudi digital retail, this versatility implies broader sociolinguistic practices of urban consumers who engage with international brands while keeping strong connections to Arabic cultural identity.

The literature reviewed reinforces the claim that localization quality is a pivotal element affecting consumer trust and involvement in Saudi online retail. As Omar et al. (2022a) and Singh and Pereira (2005) show, linguistic precision, cultural responsiveness, and functional efficiency collectively determine how users assess a website's trustworthiness. In Saudi Arabia, where internet-based purchases are still governed by trust considerations, localization breakdowns weaken brand credibility.

The findings are in alignment with confidence-centered models of digital behavior, suggested by Gefen (2000) and Kim et al. (2008). These models associate user confidence with expected credibility and identity. When website language adheres to cultural norms, it enhances a sense of belonging and moral alignment. These familiar norms may include polite Arabic expressions or appropriate gender representation. Conversely, when lexical items or visuals get in conflict with cultural norms, users may perceive the company as foreign.

Thus, localization is not a mere technical mission but a strategic discourse act that mediates between international brand identity and local user anticipations. From a professional communication perspective, this finding underscores the need to treat localization as part of corporate discourse design, not simply a translation afterthought (Li & Huan, 2019; Mahdi, 1994a, 1994b).

One of the most significant results of this review is the ongoing skills gap in professional training for localization and digital discourse. As a matter of fact, online retail in Saudi Arabia is growing fast. However, the professional development of linguistic and cross-cultural skills of practitioners has lagged behind. Bargiela-Chiappini (2014) points out that successful interaction in cross-cultural environments needs more than language proficiency. It requires understanding of discourse norms in addition to socio-pragmatic skills.

Contemporary training programs in translation and applied linguistics focus only on text-based or non-specialized translation and overlook practical skills in digital content creation and localization (Alotaibi & Aljaafari, 2024). This gap results in discrepancies across electronic retail websites and overlooked potential for linguistic innovation. Training programs do not integrate cross-disciplinary coordination between language specialists, technical developers, and brand experts. Thus, discourse techniques may be linguistically correct. However, they may overlook cultural mood or ease of use. Incorporating cross-cultural discourse theories into training programs could close this gap. This kind of integration focuses on how language forms perception, emotional resonance, and user engagement (Duizenberg, 2020; Fukuchi, 1987).

The results clearly show that intercultural competence should be regarded as a key professional skill in the Saudi e-commerce sector. This is due to the fact that intercultural competence encompasses the ability to decode and negotiate meaning across cultural contexts with awareness and accuracy (Byram, 1997). In online retail, this results in an ability to encode cultural norms in language, visuals, and user interface design.

The reviewed literature illustrates that Saudi users respond favorably to culturally respectful content. Further, they respond positively to language-sensitive content. The Saudi platforms try to generate more trust and emotional resonance by incorporating Arabic honorifics, culturally familiar words, or Islamic visuals (Ansari, 2021). Thus, intercultural training should not be limited to translation departments. It should include marketing and content design, in addition to customer service teams that deal with different audiences.

From a theoretical perspective, this aligns with Hall's (1976) and Hofstede's (2001) models, which stress the importance of understanding implicit interaction patterns and power distance in collectivist societies. Culturally adaptive electronic retail content aligns with these expectations by reinforcing relational discourse and moral congruence, both of which are important for consumer engagement in high-context environments (Breuillard, 2001; Collins & Jisum, 2019).

The presence of Arabic and English in electronic retail implies the twofold pressures of modernization and cultural conservation. English, actually, represents technological advancement, and Arabic symbolizes both national identity and cultural persistence. This bilingualism should not be regarded as a barrier. It should be viewed as a strategic discourse resource. Saudi users perceive bilingual content as both accessible and respected, as long as Arabic remains the dominant language. The optimal model, therefore, is balanced bilingualism. In other words, it is a dynamic relationship where each language performs supporting roles within online discourse (Al-Said & Al-Khalifa, 2019; Laksana, 2016; Volkov, 2019).

VI. CONCLUSION AND IMPLICATIONS

The investigation of language and multimedia content in Saudi online retail highlights that linguistic and cultural adaptation is key to the Kingdom's digital evolution. Language serves as a tool for information and also as a strategic vehicle for forming credibility and user experience. The reviewed literature repeatedly shows that localization quality, translation precision, and cross-cultural perception affect consumer perceptions of professionalism and reliability (Alharbi et al., 2023; Omar et al., 2022b). However, according to Alotaibi (2013), Danet and Herring (2007), and Ansari (2021), the irregular application of Arabic localization and the bilingual content leads to many structural barriers. It also limits consumer credibility and engagement.

Actually, professional training is essential to address all these shortcomings. Linguists, usability specialists, and marketing practitioners are in need of cross-disciplinary training. This kind of training should integrate linguistic accuracy, cultural perception, and digital discourse skills. According to Byram (1997) and Al-Said and Al-Khalifa (2019), cross-cultural proficiency is also substantial for designing culturally responsive bilingual platforms. Educators and trade stakeholders should cooperate to develop programs that ensure culturally aligned electronic content.

The presence of Arabic and English in online trade indicates that language plays two basic roles. The first role is language as a sign of cultural identity, and the second is language as a vehicle for global interaction. Both proper bilingualism and culturally appropriate content enhance affective resonance. Further, they enhance social reliability. Future research should promote interdisciplinary and data-driven methods. It is due to the fact that these methods incorporate discourse analysis and consumer analytics. In this way, future research can examine how linguistic adaptation shapes reliability, engagement, and professional practices across the Gulf region.

Building on these findings, the study offers concrete pedagogical implications for higher education and professional training in Saudi Arabia. Task-based localization projects, multimodal content adaptation workshops, and industry-aligned modules integrating translation, digital marketing, and UX writing skills can serve as practical frameworks for developing students' digital localization competencies. Educators can implement these models through curriculum

mapping that aligns course objectives with real-world digital commerce demands, structured assessment strategies that evaluate both linguistic accuracy and cultural appropriateness, and collaboration with industry stakeholders to ensure experiential learning. By adopting such evidence-based pedagogical approaches, higher education institutions can bridge the identified skills gap, foster bilingual and cross-cultural proficiency, and equip students with the tools necessary to enhance consumer trust, engagement, and professionalism in the Kingdom's rapidly evolving online marketplace.

Furthermore, integrating AI tools such as machine translation, localization QA, and generative content platforms into training can enhance students' digital localization skills. These technologies serve as pedagogical aids, supporting practical learning rather than just commercial application. Their inclusion ensures that teacher training and curriculum development remain current, industry-aligned, and responsive to evolving e-commerce demands.

ACKNOWLEDGEMENTS

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/02/36744).

REFERENCES

- [1] Addi, A., & Teimi, C. (2025). Exploring teachers' readiness to adopt an integrated language and content framework for digital design training at OFPPT institution. *Journal of English Language Teaching and Applied Linguistics*, 7(2), 190–209. <https://doi.org/10.32996/jeltal.2025.7.2.19>
- [2] Aguirre, L., & Domahidi, E. (2021). Problematic content in Spanish language comments in YouTube videos about Venezuelan refugees and migrants. *Journal of Quantitative Description: Digital Media*, 1. <https://doi.org/10.51685/jqd.2021.022>
- [3] AlGhamdi, R., Drew, S., & Alhussain, T. (2012). A conceptual framework for the promotion of trusted online retailing environment in Saudi Arabia. *International Journal of Business and Management*, 6(5), 55–57. Retrieved January 18, 2025, from <https://arxiv.org/abs/1211.2396>
- [4] Alharbi, A. S., Elyas, T., & Alghizzi, T. M. (2023). An analysis of website design and localization in Saudi Arabia: The case of Apple and Huawei. *BRU ELT Journal*, 1(1), 43–61. <https://doi.org/10.14456/bej.2023.4>
- [5] Al-Khateeb, A. (2023). Uncovering digital competence in language instruction: A study of pre-service English teachers in Saudi Arabia. *Journal of Language Teaching and Research*, 14(6), 1541–1550. <https://doi.org/10.17507/jltr.1406.07>
- [6] Alotaibi, M. B. (2013). E-commerce adoption in Saudi Arabia: An assessment of international, regional and domestic web presence. *International Journal of Information Technology and Computer Science*, 5(2), 42–56. <https://doi.org/10.5815/ijites.2013.02.05>
- [7] Alotaibi, S., & Aljaafari, M. A. H. (2024). Social commerce in Saudi Arabia: Opportunities and challenges in a digital society. *Sustainability*, 16(24), 10951. <https://doi.org/10.3390/su162410951>
- [8] Al-Said, F., & Al-Khalifa, H. (2019). Arabic–English code-switching on Saudi retail websites. *Journal of Multilingual Communication*, 11(3), 215–232.
- [9] Al-Wossabi, S. (2016). Speaking in the target language: Issues and considerations. *Journal of Language Teaching and Research*, 7(5), 886–893.
- [10] Ansari, A. (2021). Managing trust of e-customers in online shopping: A case study of Saudi Arabia. *European Journal of Business and Management*, 11(6), 67–81. <https://doi.org/10.7176/EJBM/11-6-06>
- [11] Anspoka, Z., & Eglite, I. (2019). Content and language integrated learning and teaching in digital class: Latvia experience. *Educational Role of Language Journal*, 1(1), 16–22. <https://doi.org/10.36534/erlj.2019.01.02>
- [12] Bargiela-Chiappini, F. (2014). Intercultural business discourse. In *The Routledge handbook of language and professional communication* (pp. 345–361). Routledge.
- [13] Boginskaya, O. A. (2022). Cross-disciplinary variation in metadiscourse: A corpus-based analysis of Russian-authored research article abstracts. *Training, Language and Culture*, 6(3), 55–66. <https://doi.org/10.22363/2521-442X-2022-6-3-55-66>
- [14] Casañ Pitarch, R. (2017). Language for specific purposes and graphic-adventure video games: Supporting content and language learning. *Obra Digital*, 13, 169–183. <https://doi.org/10.25029/od.2017.93.13>
- [15] Chernyavskaya, V. E., & Zharkynbekova, S. K. (2024). Code-switching patterns in Kazakh-Russian hybrid language practice: An empirical study. *Training, Language and Culture*, 8(2), 9–19. <https://doi.org/10.22363/2521-442X-2024-8-2-9-19>
- [16] Chicherina, M. M., & Chicherina, Y. V. (2021). Modern digital foreign language textbooks: Structure and content. *Nizhny Novgorod Linguistics University Bulletin*, 56, 131–142. <https://doi.org/10.47388/2072-3490/lunn2021-56-4-131-142>
- [17] Chung, L. (2012). Incorporating 3D-virtual reality into language learning. *International Journal of Digital Content Technology and its Applications*, 6(6), 249–255. <https://doi.org/10.4156/jdcta.vol6.issue6.29>
- [18] Collins, L., & Jisum, C. (2019). The role of linguistics studies on the political debate. *Linguistics and Culture Review*, 48–59. <https://doi.org/10.37028/lingcure.v3n1.15>
- [19] Contentech. (2025, March 16). *How Saudi Vision 2030 is driving content localization & digital marketing*. Retrieved May 19, 2025, from <https://contentech.com/saudi-vision-2030-content-localization-digital-marketing/>
- [20] Crystal, D. (2011). *Internet linguistics: A student guide*. Routledge.
- [21] Danet, B., & Herring, S. C. (2007). *The multilingual Internet: Language, culture, and communication online*. Oxford University Press.
- [22] Duizenberg, M. R. (2020). Linguistics landscape: A cross-culture perspective. *Linguistics and Culture Review*, 4(1). <https://doi.org/10.37028/lingcure.v4n1.17>
- [23] Esselink, B. (2000). *A practical guide to localization*. John Benjamins.

- [24] Frank, M. S., Schultz, T., & Dreyer, K. (2001). Integrating digital educational content created and stored within disparate software environments: An XML solution in real-world use. *Journal of Digital Imaging*, 14(S1), 92–97. <https://doi.org/10.1007/bf03190306>
- [25] Fukuchi, H. (1987). A localization of restrictive relative clauses. *English Linguistics*, 4, 331–335. <https://doi.org/10.9793/elsj1984.4.331>
- [26] Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, 28(6), 725–737.
- [27] Grace, R., Stratton, C., & Fonseca, F. (2019). Content matters: How online language content gives rise to digital divides. *Social Science Quarterly*, 100(6), 1999–2016. <https://doi.org/10.1111/ssqu.12691>
- [28] Hall, E. T. (1976). *Beyond culture*. Anchor Press.
- [29] Herzallah, F., Ayyash, M. M., & Ahmad, K. (2022). The impact of language on customer intentions to use e-commerce websites: Arabic consumer context. *The Journal of Asian Finance, Economics and Business*, 9(1), 273–290. Retrieved January 30, 2026, from <https://www.koreascience.kr/article/JAKO202200567709465.pdf>
- [30] Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Sage.
- [31] Ibrahim, A., Che Hussin, A. R., & Busalim, A. (2013). Framework for localization of B2C e-commerce websites in Saudi Arabia. In *2013 International Conference on Research and Innovation in Information Systems (ICRIIS 2013)* (pp. 582–587). IEEE. <https://doi.org/10.1109/ICRIIS.2013.6716774>
- [32] Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce. *Decision Support Systems*, 44(2), 544–564.
- [33] Koşar, G. (2023). An examination into pre-service English-as-a-foreign-language teachers' self-perceived technological-pedagogical-content-knowledge. *E-Learning and Digital Media*. <https://doi.org/10.1177/20427530231156169>
- [34] Korenyugina, L. M. (2025). Methods for creating visual digital educational content using Python language. *Informatics in School*, 2, 59–64. <https://doi.org/10.32517/2221-1993-2025-24-2-59-64>
- [35] Kriukova, O. (2022). Korean pop culture is reshaping Korean teaching. *ITL – International Journal of Applied Linguistics*. <https://doi.org/10.1075/itl.21010.kri>
- [36] Laksana, I. K. D. (2016). To the maintenance of Balinese culture. *International Journal of Linguistics, Literature and Culture*, 2(2). <https://doi.org/10.21744/ijllc.v2i2.122>
- [37] Lee, L. (2014). Digital news stories: Building language learners' content knowledge and speaking skills. *Foreign Language Annals*, 47(2), 338–356. <https://doi.org/10.1111/flan.12084>
- [38] Li, Y. (2025). Folk tales from diverse cultures: Digital analysis of content using natural language processing. *Applied Mathematics and Nonlinear Sciences*, 10(1). <https://doi.org/10.2478/amns-2025-0529>
- [39] Li, Z. X., & Huan, C. Y. (2019). Chinese and North American culture: A new perspective in linguistics studies. *Linguistics and Culture Review*, 3(1), 14–31. <https://doi.org/10.21744/lingcure.v3n1.13>
- [40] Lilac. (2009). Natural language processing for conceptual modeling. *International Journal of Digital Content Technology and its Applications*, 3(3). <https://doi.org/10.4156/jdcta.vol3.issue3.6>
- [41] Mahdi, W. (1994a). Some Austronesian maverick protoforms with culture-historical implications: I. *Oceanic Linguistics*, 33(1), 167. <https://doi.org/10.2307/3623005>
- [42] Mahdi, W. (1994b). Some Austronesian maverick protoforms with culture-historical implications: II. *Oceanic Linguistics*, 33(2), 431. <https://doi.org/10.2307/3623137>
- [43] Masango, C. (2006). Digital content licenses, a barrier to digital content? A South African survey. *Innovation*, 34(1). <https://doi.org/10.4314/innovation.v34i1.26533>
- [44] Mouhoubi-Messadhi, C., & Khaldi, K. (2022). Dealing with foreign language speaking anxiety: What every language teacher should know. *Training, Language and Culture*, 6(1), 20–32. <https://doi.org/10.22363/2521-442X-2022-6-1-20-32>
- [45] Müller, P., & Freudenthaler, R. (2022). Right-wing, populist, controlled by foreign powers? Topic diversification and partisanship in the content structures of German-language alternative media. *Digital Journalism*, 1–24. <https://doi.org/10.1080/21670811.2022.2058972>
- [46] Omar, A., Altohami, W. M. A., Ethelb, H., & Hamidi, B. (2022a). Localization quality assessment for more reliable e-commerce applications in Arabic. *Education Research International*, Article ID 6942735. <https://doi.org/10.1155/2022/6942735>
- [47] Omar, A., Altohami, A., & Hamidi, M. (2022b). Assessing Arabic localization quality in e-commerce applications. *Digital Communication Review*, 4(1), 52–69.
- [48] Peniro, R., & Cyntas, J. (2019). Applied linguistics theory and application. *Linguistics and Culture Review*, 3(1), 1–13. <https://doi.org/10.21744/lingcure.v3n1.7>
- [49] Rahman, M. M. (2022). The changing role of English in Bangladesh. *Training, Language and Culture*, 6(4), 20–30. <https://doi.org/10.22363/2521-442X-2022-6-4-20-30>
- [50] Rahman, S. S. (2014). Ontology localization: Bangla language. *The Dhaka University Journal of Linguistics*, 4(7–8), 27–42. <https://doi.org/10.70438/dujl/478/0002>
- [51] Shi, X., & Tian, J. (2019). On translation strategies of culture-specific items in game localization. *International Journal of Languages, Literature and Linguistics*, 5(4), 231–235. <https://doi.org/10.18178/ijlll.2019.5.4.233>
- [52] Singh, N., & Pereira, A. (2005). *The culturally customized website: Customizing websites for the global marketplace*. Elsevier.
- [53] Sun, Y.-C. (2017). Flipping every student? A case study of content-based flipped language classrooms. *E-Learning and Digital Media*, 14(1–2), 20–37. <https://doi.org/10.1177/2042753017692428>
- [54] Suvorova, L., Kravchenko, A., & Pankovyk, N. (2024). Exploring language acquisition: A person-centered approach to optimizing foreign language learning through digital content. *Grail of Science*, 35, 343–352. <https://doi.org/10.36074/grail-of-science.19.01.2024.063>
- [55] Volkov, K. (2019). Influence of culture on the language: Culture and language picture of the world. *PNRPU Linguistics and Pedagogy Bulletin*, 4. <https://doi.org/10.15593/2224-9389/2019.4.3>
- [56] von Bitter, M., & Turley, K. (2016). Deaf history, digital technology, and content-area literacy: Instructional strategies for secondary classrooms. *Sign Language Studies*, 17(1), 78–84. <https://doi.org/10.1353/sls.2016.0025>

- [57] Werner, V. (2022). Pop culture in applied linguistics. *ITL – International Journal of Applied Linguistics*. <https://doi.org/10.1075/itl.22012.wer>
- [58] Witherspoon, G. (1980). Language in culture and culture in language. *International Journal of American Linguistics*, 46(1), 1–13. <https://doi.org/10.1086/465623>
- [59] Wood, D. (2007). P. Skandera: Phraseology and culture in English [Review of the book *Phraseology and Culture in English*]. *Applied Linguistics*, 29(1), 161–163. <https://doi.org/10.1093/applin/amn003>
- [60] Woods, A. (2018). American culture: A sociological perspective. *Linguistics and Culture Review*, 2(1), 1–12. <https://doi.org/10.37028/lingcure.v2n1.6>
- [61] Xia, S. (2024). Incorporating digital multimodal composition in content teaching: A multimodal analysis of students' legal popularization videos. *Journal of Second Language Writing*, 66, 101163. <https://doi.org/10.1016/j.jslw.2024.101163>
- [62] Yousef, M., Hashem, A., Saad, H., & Hussain, K. (2012). ZLang: A scripting language for digital content creation applications. *International Journal of Computer Applications*, 50(5), 32–43. <https://doi.org/10.5120/7770-0851>

Mohammad A. Abdulaal is a faculty member at Prince Sattam bin Abdulaziz University, specializing in linguistics, applied linguistics, and theoretical linguistics. His research has garnered 131 citations, reflecting his contributions to the field. Dr. Abdulaal's research interests include dynamic assessment in EFL learning, psycholinguistic case studies, and cross-linguistic analyses of formulaic language. His work has been published in journals such as *Language Testing in Asia* and the *International Journal of English Linguistics*. A member of the academic community at Prince Sattam bin Abdulaziz University, Dr. Abdulaal contributes to the advancement of linguistic studies through both teaching and research.

Mohammad A. Aldawood is an Assistant Professor of Linguistics in the Department of English, Faculty of Arts at Port Said University, Egypt. He holds a PhD in Linguistics with a specialization in pragmatics and discourse studies. His academic interests include pragmatic theory, media discourse, and sociolinguistics in Arab and global contexts. He has authored and co-authored numerous articles in high-impact international journals on language, ideology, and communication. He has participated in several international research collaborations and academic conferences. His work aims to bridge theoretical linguistics with real-world communicative practices in media and public discourse.

From Brainstorming to Boundaries: Writing Center Tutors' Use of Generative AI

Maria Eleftheriou*

Department of English, American University of Sharjah, Sharjah, United Arab Emirates

Muhammad Ahmer

Department of Computer Science & Engineering, American University of Sharjah, Sharjah, United Arab Emirates

Abstract—The increasing presence of generative artificial intelligence (GenAI) in higher education has created new challenges and opportunities for writing centers. Peer tutors are now encountering student work that has been shaped, in whole or in part, by AI tools. Responding to this development requires balancing the pedagogical goal of supporting students' learning with institutional expectations for academic integrity. This study examines how peer tutors at the American University of Sharjah (AUS) Writing Center employ AI tools in their sessions and the boundaries they set. Data were collected from a focus group discussion, a collaboratively authored staff document, and a follow-up staff meeting. Thematic analysis of these data reveals that tutors use AI selectively for purposes such as generating ideas, clarifying language, and simplifying complex readings, while avoiding uses that would replace students' own writing or promote the use of AI. The study offers practical insights for writing centers seeking to integrate AI in ways that preserve the collaborative nature of tutoring while supporting ethical and effective student writing.

Index Terms—generative AI tools, writing centers, peer tutoring, tutor training, academic integrity

I. INTRODUCTION

Writing centers have long been spaces where students work one-on-one with peer tutors to develop their ideas, clarify their arguments, and refine their writing. This work depends on active dialogue and a shared commitment to student ownership of the writing process. At the American University of Sharjah (AUS), an English-medium university in the United Arab Emirates that serves a linguistically and culturally diverse student body, this work takes place within a multilingual context. The AUS Writing Center, modeled on the American writing center tradition, provides individualized academic writing support across disciplines through peer tutoring sessions that address a range of academic tasks, including essays, research papers, oral presentations, laboratory reports, and postgraduate theses. The center employs approximately twenty-five peer tutors and facilitates between 3,000 and 3,500 appointments each academic year, with most sessions held in person during the working week and synchronous online consultations offered on weekends.

In this setting, where peer tutors work closely with students on developing ideas and language, the arrival of generative artificial intelligence (GenAI) has disrupted established dynamics and raised new questions about authorship and originality. The speed of this change has required tutors to adjust quickly, often without established guidelines or a substantial body of research to guide their decisions. In writing center sessions at AUS, tutors are frequently among the first to engage with student writing shaped by AI tools, which require them to respond to these texts in real time while maintaining the educational value of the tutorial.

In an earlier study (Eleftheriou et al., 2025), peer tutors at AUS discussed the challenges they encountered when students brought AI-generated text to their sessions. Tutors often faced unethical requests to “humanize” AI-generated text, which were complicated by inconsistent university guidelines. To navigate this, tutors encouraged transparency about AI use to build rapport and trust, which helped them understand students' motives and tailor guidance. This openness allowed tutors to discuss the benefits of AI for brainstorming and organization while discouraging uncritical copying.

Since that study, the use of AI in higher education has expanded, and conversations have shifted from focusing solely on misconduct to exploring how AI might be incorporated constructively (Ahmad et al., 2024; Al-Alami, 2024; Al-Smadi et al., 2024; Eljack et al., 2025; Jamshed et al., 2025). However, research on AI in writing centers remains limited, particularly on how tutors integrate these tools into individual sessions without compromising the educational value of their interactions. This lack of scholarship is notable because tutors are often among the first to engage directly with student work shaped by AI. Concerns about voice, identity, and authenticity are now central, particularly for multilingual writers who move between languages and styles (Sandstead & Kibler, 2025). Tutors help students navigate both human and AI language while encouraging growth without losing individuality.

* Corresponding Author.

The present study builds on earlier findings by examining how AUS Writing Center tutors are using AI during tutorials. Drawing on data from a focus group, a collaboratively written staff document, and a follow-up staff meeting held more than a year later, the study offers an account of how tutors negotiate the role of AI in their work while supporting students' learning and meeting institutional expectations.

II. LITERATURE REVIEW

A. Challenges in Integrating GenAI in Writing Center Practice

Scholars caution that AI can displace the reasoning and reflection writing center tutorials are meant to build if students let it handle idea development, phrasing, or planning (Lingard, 2023). For multilingual writers, AI often pushes writing toward a standard academic voice, limiting translanguaging and lexical variety and distancing students from their own language resources (Barrot, 2023). This concern connects to Sandstead and Kibler's (2025) work, which shows that L2 writers' "voice," or sense of self in writing, can be lost when AI-generated phrasing replaces personal expression. They emphasize that keeping writing authentic and reflective of students' own thinking helps multilingual writers shape their voices instead of adopting AI language.

These stylistic pressures sit alongside writing center concerns about authorship and academic integrity: tutors report uncertainty about when to intervene when students present AI-shaped drafts or ask to "humanize" machine text (Roustio, 2025; Aikens & Weildon, 2025). In response, the literature points to practical solutions: keep invention and revision decisions with the student (Fitzgerald & Ianetta, 2016; Gillespie & Lerner, 2008), have students analyze AI output and then rewrite in their own voice (Lingard, 2023; Barrot, 2023), and verify any research that uses AI, since it can be wrong and fabricate citations (Khalifa & Albadawy, 2024). Our earlier study aligns with these concerns and practices: tutors described students concealing AI use, requests to "humanize" AI text, and uneven course policies. In turn, they emphasized transparency and consent in sessions and redirected work toward rewriting in the student's own language (Eleftheriou et al., 2025). These findings also reflect broader trends in higher education, where guidelines are emerging to promote transparency, AI literacy, and integrity (Moorhouse et al., 2023). Together, the scholarship suggests using AI selectively, with disclosure and verification, to support invention, reading comprehension, genre analysis, and language growth while authorship and final decisions remain with the student.

B. Opportunities Using GenAI in Writing Center Practice

Research shows that integrating AI into tutoring sessions can support students' literacy while maintaining the engagement and critical thinking central to writing center pedagogy.

(a). Brainstorming and Generating Ideas

A significant opportunity AI presents in writing is its ability to support brainstorming and idea generation. In her in-depth analysis of ChatGPT's feedback, Lingard (2023) argues that while AI has many weaknesses, such as biased, generic, and often inaccurate responses, it can serve as a practical starting point when writers begin to draft. It can suggest titles, create outlines, and help writers visualize different approaches to structuring a piece by providing multiple outline examples. Writers can then analyze and compare these options and select the most appropriate ones for their writing. Lingard recommends using "incremental prompting" to focus ChatGPT's responses and make them more helpful in brainstorming specific ideas. She illustrates this by sharing her own experience: by using ChatGPT's outline for her writing, she was able to conceptualize different sections of her paper and organize her ideas. Scholars note that AI supports prewriting and brainstorming when used critically, helping writers overcome blocks and organize ideas while tutors guide reflection (Barrot, 2023; Lingard, 2023). However, uncritical dependence on AI-generated phrasing can weaken students' reflective engagement with the writing process (Lingard, 2023; Roustio, 2025), which highlights the need for guided use. In a similar example, Krasova and Othman (2024) describe a workshop at Indiana University of Pennsylvania (IUP) where tutors identified brainstorming as the most constructive and ethical use of AI. Drawing on studies showing how students use ChatGPT to generate ideas and plan structure (Tsufim & Pomerleau, 2024) and reviews that place AI support across brainstorming, drafting, and editing (AlAfnan et al., 2023), the workshop emphasized how conversational prompting can spark thinking and help writers overcome early writing blocks. For multilingual writers, brief explanations or translations in their first language helped them clarify ideas before switching to English. Tutors encouraged students to use AI for generating possible directions, titles, questions, or outlines while relying on their own reasoning and language choices for the actual writing. They also stressed the importance of acknowledging AI assistance openly to maintain transparency and academic integrity (Krasova & Othman, 2024). This role of AI as a brainstorming partner aligns with Mahapatra's (2024) finding that ChatGPT provides formative feedback that helps students shape and refine ideas. For multilingual writers, it offers early support without replacing human judgment.

(b). Assisting With Research

In addition to brainstorming at the topic and outline stage, AI can help students with research and writing. Researchers have discussed ways in which AI tools can assist students in the research process by facilitating research tasks (Khalifa & Albadawy, 2024; Ringo, 2025; Larios Soldevilla et al., 2025). Khalifa and Albadawy (2024) claim that

AI brainstorming can lead writers to innovative research questions by identifying gaps in the literature and steering their research in new directions. Additionally, they argue that AI can assist students in organizing literature reviews by suggesting logical structures and helping outline the flow of ideas to ensure clarity and cohesion. It can also reduce the time students spend searching for relevant articles by efficiently summarizing them; this approach enables quicker engagement with large amounts of academic literature. However, scholars consistently caution that human oversight remains essential, since AI tools can misrepresent sources and fabricate citations (Khalifa & Albadawy, 2024). This idea aligns with growing evidence that tutors and instructors can use AI tools to support research-related literacy while emphasizing the importance of human guidance (Moorhouse et al., 2023). As university guidelines increasingly acknowledge AI's role in academic writing, ethical integration rather than restriction is becoming a key focus in tutor education.

(c). Supporting Multilingual Writers Through AI Feedback

AI tools can play a critical role in developing multilingual students' writing skills. The interactive space of ChatGPT has positively impacted multilingual learners by improving their reading, writing, and vocabulary skills (Monika & Suganthan, 2024). Barrot (2023) explains that by offering automatic, specific feedback on written work, AI can guide multilingual students on issues they typically struggle with, such as coherence, cohesion, and text organization. AI allows tutors to focus on higher-order aspects like argumentation and structure. This adaptive support, as Mahapatra (2024) observes, helps learners refine syntactic and lexical skills while receiving context-specific feedback. Researchers have also found that AI is particularly effective for vocabulary acquisition. Monika and Suganthan (2024) found that ChatGPT was effective for vocabulary acquisition when ESL learners were exposed to diverse, meaningful contexts. This exposure allows students to encounter and practice new words naturally, supporting long-term retention. These findings reflect Abdelhalim and Alsehibany's (2025) study with Saudi EFL learners, which demonstrated that ChatGPT-based vocabulary scaffolding significantly improved productive vocabulary and confidence in word choice. Their participants emphasized that AI's interactive correction and context-rich prompts made language practice both more engaging and more reflective of honest communication.

Similarly, Alharbi and Khalil (2024) emphasize that ChatGPT's real-time feedback enables students to make and learn from errors on the spot. The interactive nature of these tools also makes vocabulary learning more engaging, encouraging consistent practice. Recent research in writing center studies expands this perspective. Lundin (2025) found that tutors notice multilingual writers using ChatGPT to bridge gaps in language ability, observing that "a student wants to see how a sentence might be reframed, and the reframing of that sentence is beyond their linguistic capability at that moment." This process enables writers to test and expand their linguistic repertoire much as they would when encountering new vocabulary through exposure and experimentation. Nevertheless, Lundin cautions that overreliance on AI may reinforce standardized or "colonial" forms of English, which demonstrate an ongoing tension between linguistic empowerment and conformity. These concerns echo Sandstead and Kibler's (2025) advice for instructors to help multilingual students retain authorial agency and cultural identity in AI-assisted writing, encouraging them to negotiate their authentic linguistic voices rather than suppress them.

(d). Toward Ethical and Reflective AI Use in Writing Centers

While concerns about AI's potential to reduce students' engagement with writing are valid, a growing body of research highlights its pedagogical promise when used critically and transparently. ChatGPT, for instance, can "support bottom-up writing skills, freeing up time, space, and energy for more advanced aspects of composition" (Daniel et al., 2023, p. 37). Mahapatra (2024) explains that when ChatGPT is used as a form of dialogic feedback, it encourages reflection, self-monitoring, and independence in learning. This approach complements the peer and tutor feedback models that are already central to writing center practice. Mollick and Mollick (2023) describe multiple ways students can use AI as a writing partner, from receiving real-time feedback to refining style and generating creative prompts that encourage fluency. These studies, along with initiatives such as the IUP workshop described by Krasova and Othman (2024), highlight a growing recognition that AI can serve as a productive and ethical aid in writing instruction when used to support learning rather than replace it. Their professional development model demonstrated how tutors can analyze both human and AI feedback, discuss challenges collaboratively, and co-create policy frameworks that emphasize transparency and student ownership of writing. Despite these promising developments, little is known about how peer tutors themselves apply such principles during real tutorials, a gap this study addresses.

III. METHODOLOGY

This section describes the research design and outlines the study context, participants, data collection methods, and analysis procedures.

A. Research Question

We address the following research question: To what extent do writing center tutors incorporate AI tools into tutoring sessions?

B. Participants

The participants were drawn from the peer tutoring staff of the AUS Writing Center. Peer tutors at the center are generally bilingual and demonstrate high academic achievement. Most are recruited from advanced undergraduate writing composition courses, while others are graduate students enrolled in the Master of Arts in Teaching English to Speakers of Other Languages program. All tutors undertake a semester-long, credit-bearing training course before beginning their work. This training equips them with the practical skills and pedagogical knowledge needed to support students effectively in one-on-one sessions.

C. Instruments

The three instruments used are the "Focused Group Discussion", "Collaborative Staff Document", and "Transcript of Fall 2024 Staff Meeting". All materials underwent reflexive thematic analysis (Braun & Clarke, 2006). The datasets were read multiple times to ensure familiarity with the content. Coding proceeded inductively, with initial codes generated from meaningful units of text and iteratively refined through constant comparison across sources. Codes were then organized into themes, which were reviewed against the coded extracts and the complete data set before being defined and named. Coding, memoing, codebook versioning, and data retrievals were managed in ATLAS.ti, including use of code–co-occurrence queries and memo links to support transparent theme development.

The two researchers checked the codes and themes and resolved differences through discussion. We kept dated notes during the analysis to show how our ideas and themes developed over time. We triangulated across data sources and across time. We also examined counterexamples to test and refine the themes.

D. Data Collection Methods

Three qualitative data sources were used to answer our question: a focus group discussion, a collaboratively authored staff document, and a transcript of a Fall 2024 staff meeting. Using multiple data sources enhanced triangulation and credibility.

(a). Research Design

Approval of the research was granted by the Institutional Review Board (IRB). All procedures complied with IRB requirements for informed consent, confidentiality, and secure data handling. Purposive sampling was then used to identify tutors with experience with GenAI in their tutorials who wished to participate in the study. This strategy ensured information-rich cases aligned with the research aim (Patton, 2015). An initial survey was distributed to the whole tutoring staff. The survey asked whether tutors had worked with students who had used AI and whether they were interested in joining a focus group to discuss this experience. Ten tutors participated: this number is consistent with recommendations in the literature that suggest a group of six to ten participants provides a balance between diversity of viewpoints and the opportunity for each person to contribute meaningfully (Freitas et al., 1998). The final group reflected variation in academic major, tutoring experience, and linguistic backgrounds to maximize heterogeneity and perspectives (Table 1).

In addition, two naturally occurring professional-development data sources were included: a spring 2024 workshop that produced a collaborative document in which tutors jointly outlined strategies and guidelines for responding to AI in student work and listed the tools they incorporated into sessions; and a Fall 2024 follow-up meeting that generated a transcript capturing shifts in tutors' perceptions, their strategies for addressing AI-related issues, and their discussion of how they use AI tools in sessions.

(b). Focus Group Discussion

The focus group was held in March 2024 in a quiet room within the writing center and lasted two hours. Ten tutors participated after providing written consent. Participants received an information sheet explaining the study's aims, the voluntary nature of participation, and the right to withdraw at any time without penalty. To protect confidentiality, pseudonyms were assigned to each participant, and all data were anonymized. A semi-structured interview protocol guided the discussion, balancing consistency across core questions with flexibility to probe emergent issues. Questions invited tutors to describe challenges they faced when students used AI, outline strategies they found compelling, and reflect on ethical concerns. The conversation was audio-recorded in full and transcribed verbatim for analysis. Field notes captured key gestures and context to deepen understanding. Participants for the focus group discussion are listed in Table 1.

TABLE 1
DEMOGRAPHIC INFORMATION FOR THE FOCUSED GROUP DISCUSSION PARTICIPANTS

Variables		N	Percentage
Major	Architecture	2	20
	Computer Science	2	20
	English	2	20
	Industrial Engineering	2	20
	Journalism	1	10
	Psychology	1	10
	Total	10	100
Standing	Junior	6	60
	Senior	3	30
	Fifth Year	1	10
	Total	10	100
Writing Center Semesters	1	1	10
	2	4	40
	3	1	10
	4	3	30
	4 >	1	10
	Total	10	100
Sex	Female	9	90
	Male	1	10
	Total	10	100
Nationality	Egyptian	4	40
	Indian	1	10
	Lebanese	1	10
	Palestinian	1	10
	Pakistani	1	10
	Sudanese	1	10
	Tunisian	1	10
	Total	10	100

(c). *Collaborative Staff Document*

In April 2024, approximately one month after the focus group, we held a staff meeting in which tutors contributed ideas to a shared Google document. The activity was a 60-minute structured professional-development exercise with a brief orientation, two contribution rounds, and a final synthesis round. Its purpose was to gather practical suggestions for handling common AI-related scenarios in tutoring sessions. Prompts were derived from the focus group's preliminary codes and then reviewed by the two researchers for clarity and alignment with the study aims. Prompts asked tutors, for example, how they might respond if a student concealed AI use or requested that AI-generated text be revised to appear more human, and they also invited tutors to list the AI tools they used in sessions and to comment on the observed impact of those tools. The format allowed participants to add ideas in real time, build on one another's responses, and record a broad range of strategies. Access to the document was restricted to attending tutors and the research team, and all entries in the research corpus were de-identified and imported into ATLAS.ti as a primary text; individual entries and comment threads were treated as meaning units and linked to analytic memos for transparency. Portions of this dataset were analyzed in a prior publication (Eleftheriou et al., 2025). The present paper extends that work by focusing on tutors' accounts of the tools they incorporated into sessions and the perceived effects on students' research and writing. The document also served as a naturally occurring institutional text for triangulation and for time-based comparison with the Fall 2024 staff-meeting transcript. Details about the tutors who contributed to the collaborative discussion and document appear in Table 2.

TABLE 2
DEMOGRAPHIC INFORMATION FOR THE COLLABORATIVE DOCUMENT GROUP PARTICIPANTS

Variables		N	Percentage
Major	Biology	2	9.5
	Chemical Engineering	1	4.8
	Computer Engineering	1	4.8
	Computer Science	1	4.8
	Economics	1	4.8
	English	4	19
	Environmental Science	1	4.8
	Industrial Engineering	4	19
	International Studies	3	14.2
	Journalism	1	4.8
	Psychology	2	9.5
	Total	21	100
Standing	Junior	12	57.1
	Senior	9	42.9
	Total	21	100
Writing Center Semesters	1	6	28.6
	2	5	23.8
	3	3	14.3
	4	2	9.5
	4 >	5	23.8
	Total	21	100
Sex	Female	19	90.5
	Male	2	9.5
	Total	21	100
Nationality	Egyptian	4	19.0
	Indian	5	23.8
	Lebanese	2	9.5
	Pakistani	2	9.5
	Palestinian	5	23.8
	Sudanese	1	4.8
	Tunisian	1	4.8
	Turkish	1	4.8
	Total	21	100

(d). Transcript of Fall 2024 Staff Meeting

The Fall 2024 meeting was designed to revisit AI in tutoring sessions after more than a year of institutional and technological change. It served as a follow-up to the Spring 2024 focus group and collaborative document. The discussion brought together members of the tutoring staff and explored whether earlier strategies remained effective, how faculty policies had shifted, and what new challenges had emerged. The meeting followed a semi-structured agenda circulated in advance, and the director facilitated the discussion. It is important to note that not all staff participated in this meeting. Because attendance was partial, findings from this meeting reflect the perspectives of those present. Triangulation with the Spring 2024 focus group and the collaborative document helps mitigate this constraint.

The session lasted sixty minutes and was conducted online via Google Meet, a university-approved platform. With participants' consent, the discussion was audio-recorded and later transcribed verbatim. The research team verified transcription accuracy.

The transcript was uploaded to ATLAS.ti and analyzed alongside the earlier focus group and the collaborative staff document. Cross-source comparison supported data triangulation and allowed identification of both recurring and newly emergent themes. This additional data source made it possible to trace developments over time, revealing areas of continuity, adaptation, and change in tutors' perceptions, ethical reasoning, and practices related to GenAI. Demographic information for the staff meeting is presented in Table 3.

TABLE 3
DEMOGRAPHIC INFORMATION FOR THE STAFF MEETING PARTICIPANTS

Variables		N	Percentage
Major	Architecture	1	10
	Biology	1	10
	Computer Science	1	10
	Environmental Science	1	10
	Industrial Engineering	2	20
	Journalism	1	10
	Physics	1	10
	Psychology	2	20
	Total	10	100
Standing	Junior	3	30
	Senior	7	70
	Total	10	100
Writing Center Semesters	1	4	40
	2	2	20
	3	2	20
	4	1	10
	4 >	1	10
	Total	10	100
Sex	Female	10	10
	Male	0	0
	Total	10	100
Nationality	Egyptian	2	20
	Indian	4	40
	Lebanese	1	10
	Saudi Arabian	1	10
	Sudanese	1	10
	Turkish	1	10
	Total	10	100

IV. RESULTS AND DISCUSSION

A. Using AI as a Tool in Sessions

The section below explores the response to our research question, “To what extent do writing center tutors incorporate AI tools into tutoring sessions?”

Educators have cautioned against using GenAI tools in the writing process. However, as Stowe (2023), a writing center director, writes, “As long as tutors are very careful about checking the information they find, these programs could be useful when helping students with research. They work, in some ways, as quick search engines, and as long as you check the information you find, that might be a place to use them in a practical way in our centers” (para. 2). Our findings align with his idea. Tutors revealed the various ways they integrate AI into their sessions across all datasets: they incorporate GenAI with caution, promoting its use for acceptable purposes while ensuring it complements, rather than replaces, student learning. Similarly, Mollick and Mollick (2023) argue for structured exploration that helps tutors evaluate AI’s usefulness while maintaining human agency. These shifts suggest that tutors are becoming innovators, but their work must remain guided by pedagogical and ethical principles.

Comparable developments are evident in the workshop by Krasova and Othman (2024) at the Kathleen Jones White Writing Center. Their professional development initiative, “Overview of AI Technology and its Relevance to Writing Center Support,” invited tutors to examine benefits and risks, try tools in a guided way, and compare ChatGPT feedback with human feedback on the same student paper. Tutors noted that AI feedback tended to focus on surface-level edits, while human feedback supported rhetorical development and writer confidence. The closing activity asked tutors to draft a shared AI policy that encouraged transparent use of AI while preserving student agency. This structured reflection helped tutors clarify ethical boundaries in ways that mirror our participants’ priorities.

By comparison, AUS tutors reached similar conclusions through collaborative discussion and reflection on real tutoring experiences. Both contexts show that when tutors engage critically with AI, they arrive at a shared position: AI can help with brainstorming, clarifying ideas, and modeling feedback practices, but it cannot replace real conversation and connection in tutoring. This finding also aligns with Mahapatra’s (2024) argument that ChatGPT functions as a dialogic feedback tool that helps students reflect and self-regulate during writing.

B. Brainstorming and Generating Ideas

In the focus group discussion, tutors described using AI for brainstorming; they found it helpful for stimulating discussion. Additionally, it proved valuable for suggesting titles and initiating conversations about effective choices. As Roha explained during the focus group discussion, many students end sessions unsure how to title their essays, so she tells them, “Ask ChatGPT—it’s amazing at titles. Don’t waste your time on that; focus on the writing first, and if you’re stuck at the end, you can ask ChatGPT for a title.” This approach shows tutors see AI as useful for minor tasks while

focusing on substantive writing. Some tutors used ChatGPT to prepare for sessions on unfamiliar topics. For example, in the focus group discussion, Sahar said that ChatGPT helped her prepare for a session with an engineering student because she lacked knowledge in the field, and she felt better prepared after learning from it. Similarly, tutors mentioned brainstorming in the collaborative document. Jamila explained, "I would encourage them to implement some of the ideas they received from ChatGPT but use their own voice." This highlights how tutors encourage students to use AI as a creativity booster while ensuring that the student's original voice remains central to the writing process. Tutors at other writing centers have also started incorporating AI into their sessions: Tutors at the University of Connecticut experimented with AI and demonstrated how it can assist in unblocking writers, refining thesis statements, and generating conclusions (Deans et al., 2023). In these cases, the tutors emphasized the importance of using AI as a complement to human input, rather than a replacement for it. At the same time, in the transcript of our Fall 2024 staff meeting, Roha and Jamila expressed reservations about AI-supported brainstorming, noting that it can omit the exploratory talk and free-writing that typically lead to stronger ownership of ideas. They indicated a preference for traditional brainstorming (paper lists, clustering, tutor-led questioning) and described it as more creative and a better model of the thinking process. For these tutors, the work of generating, sorting, and revising ideas with students felt more authentic and transferable than prompting an AI tool. Their views align with research emphasizing the role of tutor modeling and back-and-forth conversation in building genuine idea development (Fitzgerald & Ianetta, 2016) and with findings that students often value the spontaneity and co-construction of traditional brainstorming over computer-generated ideas (Kim et al., 2024). Similarly, Mollick and Mollick (2023) argue that while AI can serve as a creative partner, its results may narrow imaginative exploration when over-relied upon, making human collaboration and reflective questioning crucial for deep learning. Studies by Barrot (2023) and Krasova and Othman (2024) note that while AI can efficiently generate prompts, authentic creativity and rhetorical awareness still emerge most strongly through tutor guidance and human conversation.

C. Assisting With Research

Data from both the focus group discussion and the collaborative document revealed that tutors use AI as a scaffold for academic research to simplify complex scholarly sources and assist with keyword generation. As tutors noted in the collaborative document, AI can "simplify a scholarly paper into simpler words for them to understand," allowing students to access the key ideas without being confused by dense language. Sana expanded on this point in the same document by observing that "you can use ChatGPT to help explain sources if they are written in complicated words or jargon. You can copy paste some paragraphs... to help explain it in simpler terms." By transforming complex language into summaries, tutors help students build confidence to engage with academic texts and reduce the cognitive load that often accompanies difficult readings. However, this form of support still requires students to engage directly with sources to interpret, evaluate, and accurately cite scholarly ideas. This finding aligns with Khalifa and Albadawy's (2024) description of AI's ability to support comprehension and content development by clarifying specialized academic language and providing accessible paraphrases that serve as entry points to deeper learning. It also connects with Mahapatra (2024), who views AI as a formative feedback partner that supports reflection, and with Sandstead and Kibler (2025), who emphasize the need to preserve authenticity and voice even when AI simplifies complex academic material.

Tutors also identified AI's usefulness in generating search terms and noted that it can help students identify relevant keywords and refine their research focus. As Jamila said in the focus group discussion, "ChatGPT is great for coming up with search terms when students are stuck. It helps them find the right direction when doing research." She also described the collaboration this process creates: "We went to ChatGPT and asked for a few keywords, and what I found is that it enhances the session because you work more collaboratively—I just thought that it's kinda like teamwork." In this way, AI serves as a shared prompt that both the tutor and the student examine together. It gives them a starting point for discussion, lets them test and refine terms side by side, and keeps decision-making visible. This collaborative use aligns with the writing center's mission to support learning through conversation while keeping ownership with the student. This use of AI aligns with Kim et al.'s (2024) observation that generative tools can facilitate research discovery by enhancing students' ability to navigate databases and identify conceptual links between topics. Through these practices, tutors are not replacing critical inquiry but showing how AI can support the early stages of research from conceptualization to information gathering.

D. Assisting With Grammar

Data from the focus group discussion and collaborative document showed that tutors promote the careful use of ChatGPT for grammar correction and phrasing improvement. While this was not as central as other forms of AI use, tutors recognized its pedagogical value when integrated thoughtfully. As Samira explained, "I told her, 'You have to use it in a way that helps you build the skill because like 'teach me how to fish, don't give me fish.''" This reflective stance ensures that students learn from the revision process rather than relying on AI as a shortcut. For multilingual students at AUS, grammatical accuracy often represents a persistent challenge, and tutors observed that AI can help students recognize and self-correct errors through immediate, example-based feedback. This aligns with research showing that AI-generated feedback can enhance language awareness by making grammatical structures more visible to learners and improving syntactic fluency (Khalifa & Albadawy, 2024; Kim et al., 2024). Tutors emphasized that this

support must contribute to deeper learning, not only surface-level correctness. This is consistent with research showing that AI grammar tools can strengthen accuracy and promote learner autonomy, but only when students actively engage with feedback rather than apply changes without reflection (Mohebbi, 2025).

In the collaborative discussion, Adam explained that decisions about AI use must be tied to course learning outcomes. He explained that “it is possible to connect contexts to the learning outcomes and engage in reflection on whether particular GPT use violates those learning outcomes. For example, using Grammarly for WRI 001 [beginner-level writing course] is not okay, but for ENG 204 [advanced writing course] it is. The way we identify is based on what the purpose of the course is.” His insight demonstrates tutors’ ethical awareness of developmental differences and of the assignment’s purpose when shaping AI-supported writing. Ultimately, tutors positioned AI grammar tools as a formative support, a starting point for discussion rather than a final product. This approach ensures that students remain active participants in their linguistic development, which is a core principle of writing center pedagogy and ethical AI integration.

E. Generating Model Texts

Across the collaborative document, focus group, and staff meeting, tutors described using ChatGPT to generate model texts that students could learn from rather than replicate. In the focus group discussion, Samira recounted asking ChatGPT to produce an introduction for a genre of essay on a different topic, which helped a struggling student understand how introductions are typically structured. Once the model text was generated, she told her student, “Okay, now that you figured it out from ChatGPT, try applying the same concepts on your own ideas.” This example shows how tutors frame AI output as a learning scaffold, a temporary support that helps students develop independence in their writing.

In the collaborative document, Samira described a guided approach to using AI models: “Learning from AI, if there is a particular aspect they like about the generated text, we analyze it together and learn to incorporate it intentionally into original writing.” This practice treats AI output as a scaffold students study and then rewrite in their own voice. This practice is consistent with Mahapatra’s (2024) argument that structured feedback helps students reflect on and internalize patterns rather than imitate them (Mahapatra, 2024). By stressing analysis and tailoring to the student, tutors do what Roustio (2025) describes: they shape the conversation, so students use AI responsibly and keep their agency.

The staff meeting data further reinforced this trend. Roha explained, “Instead of telling them to start from scratch, I tell them, ‘Find one general [AI-generated] paragraph that you like.’ And then we’d walk through it to outline the major points in the paragraph—for example, the topic sentence, the source, the explanation.” Hania agreed and described a similar practice: “I take the model paragraph and have students try to structure their sentences like that.” These practices align with Khalifa and Albadawy (2024), who found that GenAI supports content development by illustrating clear organizational models that students can emulate. Likewise, Kim et al. (2024) show that such guided engagement with AI models promotes metacognitive awareness: students begin to reflect on how texts are constructed and what makes them effective.

F. Assisting With Vocabulary

Tutors described using ChatGPT for quick lexical support during sessions, especially when students could not find the right word in English. This pattern appeared across the focus group, the collaborative document, and the staff meeting, and it is relevant in a multilingual setting like AUS. In the focus group, Layla discussed the limits of the process and the importance of student ownership. However, she allowed targeted language help: “Most of the processes of writing should be off-limits, maybe brainstorming or vocabulary retention (what’s the word for...?), I would even be okay with sentence generation, so long as they rephrase it before adding it to their essay.” In the collaborative document, Sana took a similar position, using AI as a helper rather than a writer: “I’d probably use a lot of the same approaches of trying to promote original writing. I think I’d also maybe try to point out ways in which ChatGPT could be helpful like [with]...vocabulary.” Tutors emphasized that this use should remain interactive and critical. Their approach parallels Abdelhalim and Alsehibany’s (2025) findings that AI-supported vocabulary practice can expand lexical range and confidence when used reflectively. During the staff meeting, tutors reported asking for lexical assistance that keeps ownership with the student. Noora explained that she tells students, “You can use ChatGPT in ...scenarios where you want to ask for it for like transition words, or maybe like, just asking if the sentence is flowing well...but just don’t trust this input all the time, because it doesn’t give you the right answers every time.” Across all datasets, tutors use AI to prompt vocabulary and keyword possibilities, then ask students to rephrase and integrate choices in their own voice. This keeps the efficiency benefits of AI while supporting language development, originality, and clarity for multilingual writers.

These accounts show a consistent approach: tutors use AI to expand vocabulary in the moment, then return the writing to the student through guided discussion and rewriting. This aligns with research that treats AI text as material for analysis rather than text to adopt, a stance that helps protect multilingual voice and invites metalinguistic reflection (Barrot, 2023; Lingard, 2023). Studies with multilingual learners report that immediate, example-based feedback can support vocabulary growth and confidence when students review suggestions critically and then rephrase in their own words (Monika & Suganthan, 2024; Kim et al., 2024). When AI suggests vocabulary or simplifies a challenging passage, tutors can model verification and source-aware phrasing so students do not copy formulations that may be

inaccurate or out of context (Khalifa & Albada, 2024). This practice is also consistent with writing center principles that prioritize student ownership and collaborative conversation, keeping invention and revision decisions in the student's hands (Fitzgerald & Ianetta, 2016; Gillespie & Lerner, 2008), and with multilingual pedagogy that values movement across languages to develop precise academic vocabulary (Zhang & Huang, 2024).

G. Modeling and Consent: Navigating Ethical Boundaries in AI Use

A consistent theme in the staff meeting, a year after our focus group discussion, was tutors' apprehension about modeling the use of GenAI during sessions. Several worried that visible use could normalize or encourage it. Banaa explained, "I haven't used or encouraged ChatGPT because, like what Hania said, I'm scared it makes them think they can use it for writing the whole paper." Roha agreed, "Let's say we want to learn how to write an introduction using a ChatGPT-generated paragraph. I feel like it would be an incentive for students to use ChatGPT after the session. I feel like I'd be promoting its use for their introduction." Jamila added, "Because we cannot monitor AI use, we will not know the extent of how they are using it or whether they are using it correctly".

As a result, many tutors resisted presenting ChatGPT as a publicly endorsed tool in the writing center. Noora even concealed her own occasional use to avoid signaling approval: "ChatGPT gives me a list of transition words. I read it as if it is not from ChatGPT." Her acknowledgment of its usefulness, paired with deliberate concealment, captures a central tension in the data. Tutors want to draw on helpful resources without encouraging dependency or undermining student ownership.

Equally important, tutors stressed that any in-session use of AI should be limited to the student's explicit consent. They viewed transparency and consent as essential safeguards that protect the collaborative nature of tutoring and preserve the student's agency in the learning process. This emphasis parallels writing center scholarship that situates ethical tutoring in practices of openness, reciprocity, and shared decision-making (Gillespie & Lerner, 2008; Fitzgerald & Ianetta, 2016). In the context of AI, recent studies have similarly stressed that transparency and consent are key to maintaining equitable and ethical engagement (Aikens & Weildon, 2025; Roustio, 2025). These values also align with broader ethical frameworks for AI in education that highlight human agency, accountability, and informed participation as foundations of responsible use (Floridi & Cowls, 2021). When tutors decide to demonstrate an AI tool, they should first explain its purpose, clarify how it will be used, and ensure the student agrees. This practice situates AI as a shared pedagogical choice rather than a unilateral decision made by the tutor.

Reflexive note. Although not part of the formal dataset, unsolicited student feedback to the writing center survey after the study echoed these concerns. One student reported that a tutor "simply copied everything into an AI tool to find answers," adding that they "showed no interest in reading or understanding the material herself," and described the session as "a frustrating waste of time." While anecdotal, this feedback emphasizes the importance of student consent and transparency when AI is used in real-time. Visible, uninvited use can compromise trust and the perceived authenticity of the session, reinforcing tutors' view that AI should only be introduced through open discussion and mutual agreement.

V. CONCLUSION

A. Findings

This study contributes to the growing scholarship on GenAI in writing centers by examining how tutors at the AUS Writing Center integrate these tools while maintaining student-centered learning. Across all datasets, tutors described AI as a valuable support for brainstorming, language development, and research, while emphasizing that human interaction remains central to the tutorial. Their selective use of ChatGPT reflects a careful balance between pedagogy, ethics, and transparency.

These findings show that tutors are not simply responding to technological change but actively shaping responsible practices for AI use. Through reflection and collaboration, they position themselves as guides who help students navigate new writing technologies. More broadly, the study shows that core writing center values such as collaboration, reflection, and student ownership continue to guide tutoring practice even as tools evolve. When used thoughtfully and transparently, AI can support these principles.

B. Implications

These findings have several implications for writing center administration and tutor training. Tutor education can include AI literacy and ethical reasoning so that tutors can evaluate when AI use supports learning. Writing centers can also develop clear guidelines for AI use that align with university policies while still allowing tutors room for professional judgment. Further, given the linguistic diversity of students at AUS and comparable institutions, training should include strategies for using AI to support multilingual writers without replacing meaningful language learning.

Finally, this study contributes a regional perspective to international discussions about writing, authorship, and the role of AI in higher education. Documenting tutor decision-making in a multilingual Middle Eastern context highlights the importance of including diverse experiences in shaping the future of writing center work in the age of GenAI.

C. Recommendations for Future Research

As a single-site qualitative study, the findings reported in this paper are not intended to be generalizable. The perspectives represented reflect the experiences of peer tutors at one English-medium institution, and the data draw primarily on tutors' self-reported accounts rather than direct observation of live tutorials. These limitations point to several directions for future research.

Future studies could complement tutor perspectives with student input to further validate and contextualize findings related to AI-integrated writing center sessions. Incorporating student interviews, surveys, or reflective feedback would allow us to examine how students perceive the usefulness of AI-supported tutoring, how they understand the boundaries between human and AI contributions, and whether these boundaries align with tutors' intentions.

Additionally, observational approaches can be used to examine how AI use is negotiated during actual tutoring sessions. Observing tutor–student interactions would enable researchers to analyze how decisions regarding transparency, consent, and the appropriate use of AI unfold in real-time, rather than relying solely on retrospective accounts. Taken together, triangulating tutor perspectives with student experiences and observational data would strengthen the technical soundness of future research and offer a more holistic understanding of how AI is integrated into writing center practice.

REFERENCES

- [1] Abdelhalim, S. M., & Alsehibany, R. (2025). Integrating ChatGPT for vocabulary learning and retention: A classroom-based study of Saudi EFL learners. *Language Learning & Technology*, 29(1), 1–24. <https://doi.org/10.64152/10125/73635>
- [2] Ahmad, D., Latif, I., Arafah, B., & Suryadi, R. (2024). Defining the role of artificial intelligence in improving English writing skills among Indonesian students. *Journal of Language Teaching and Research*, 15(2), 568–578. <https://doi.org/10.17507/jltr.1502.25>
- [3] Aikens, K., & Weildon, H. (2025). Generative AI tutor education in our writing center: A slow approach. *The Peer Review*. Retrieved August 14, 2025, from <https://thepeerreview-iwca.org/issues/issue-9-1/generative-ai-tutor-education-in-our-writing-center-a-slow-approach/>
- [4] AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3(2), 60–68. <https://doi.org/10.37965/jait.2023.0184>
- [5] Al-Alami, S. E. (2024). EFL Learners' Attitudes Towards Utilizing ChatGPT for Acquiring Writing Skills in Higher Education: A Case Study of Computing Students. *Journal of Language Teaching & Research*, 15(4). <https://doi.org/10.17507/jltr.1504.01>
- [6] Alharbi, K., & Khalil, L. (2024). Artificial Intelligence (AI) In ESL Vocabulary Learning: An Exploratory Study on Students and Teachers' Perspectives. *Migration Letters*, 20(S12), 1030–1045. <https://migrationletters.com/index.php/ml/article/view/8224>
- [7] Al-Smadi, O. A., Ab Rashid, R., Saad, H., Zrekat, Y. H., Kamal, S. S. L. A., & Uktamovich, G. I. (2024). Artificial intelligence for English language learning and teaching: Advancing sustainable development goals. *Journal of Language Teaching and Research*, 15(6), 1835–1844. <https://doi.org/10.17507/jltr.1506.09>
- [8] Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- [9] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- [10] Daniel, S., Pacheco, M., Smith, B., Burriss, S., & Hundley, M. (2023). Cultivating writerly virtues: Critical human elements of multimodal writing in the age of artificial intelligence. *Journal of Adolescent & Adult Literacy*, 67, 32–38. <https://doi.org/10.1002/jaal.1298>
- [11] Deans, T., Praver, N., & Solod, A. (2023). *AI in the writing center: Small steps and scenarios*. University of Connecticut. Retrieved August 18, 2025, from <https://dept.writing.wisc.edu/blog/ai-wc/>
- [12] Eleftheriou, M., Ahmer, M., & Fredrick, D. (2025). Balancing ethics and support: Peer tutors' experiences with AI tools in student writing. *Contemporary Educational Technology*, 17(3), ep587. <https://doi.org/10.30935/cedtech/16554>
- [13] Eljack, N. S., Altahir, N. A., Mohammed, N. I., & Osman, S. A. (2025). ChatGPT as an AI Assistant Tool to Enhance University Students' Teaching, Assessment and Academic Achievements: A Case Study of an EAP Lecturer. *Journal of Language Teaching and Research*, 16(6), 1813–1823. <https://doi.org/10.17507/jltr.1606.15>
- [14] Fitzgerald, L., & Ianetta, M. (2016). *The Oxford guide for writing tutors: Practice and research*. Oxford University Press.
- [15] Floridi, L., & Cowls, J. (2021). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 3(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- [16] Freitas, H., Oliveira, M., Jenkins, M., & Popjoy, O. (1998, February). *The focus group: A qualitative research method* (Working Paper No. ISRC 010298). Information Systems Research Centre, Merrick School of Business, University of Baltimore. Retrieved July 22, 2025, from <https://www.scirp.org/reference/referencespapers?referenceid=2657165>
- [17] Gillespie, P., & Lerner, N. (2008). *The Longman guide to peer tutoring* (2nd ed.). Pearson Longman.
- [18] Jamshed, M., Albedah, F., Hussain, N., & Banu, S. (2025). Assessing the Empowering Efficacy and Perceived Pedagogical Prospects of ChatGPT in the Saudi Arabian Context: A Mixed Method Study. *Journal of Language Teaching and Research*, 16(3), 1002–1012. <https://doi.org/10.17507/jltr.1603.31>
- [19] Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, 5, 100145. <https://doi.org/10.1016/j.cmpbup.2024.100145>
- [20] Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12878-7>

- [21] Krasova, A., & Othman, M. (2024). Writing Center Instruction for the Age of AI: Tutors Professional Development Workshop. *The Peer Review*, 9(2). Retrieved October 29, 2025, from <https://thepeerreview-iwca.org/issue-9-2/writing-center-instruction-for-the-age-of-ai-tutors-professional-development-workshop/>
- [22] Larios Soldevilla, O. A., Mendoza Ibarra, V., Urdanegui Sibina, R., Zilberman, J., Lizola-Margolis, P. E., Flores Peraltilla, A. V., Ghersi, J. H., Valencia Brenis, M., & Aviles Valdez, J. R. (2025). Adoption of AI tools and their impact on the academic research output of business school students. *Cogent Education*, 12(1), Article 2556892. <https://doi.org/10.1080/2331186X.2025.2556892>
- [23] Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. *Perspect Med Educ*, 12(1), 261–270. <https://doi.org/10.5334/pme.1072>
- [24] Lundin, I. (2025). Emotional nuance and tension present in writing center tutors' perceptions of AI. *The Peer Review*, 9(2). Retrieved June 3, 2025, from <https://thepeerreview-iwca.org/issue-9-2/emotional-nuance-and-tension-present-in-writing-center-tutors-perceptions-of-ai/>
- [25] Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 9. <https://doi.org/10.1186/s40561-024-00295-9>
- [26] Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: a systematic review. *Cogent Education*, 12(1). <https://doi.org/10.1080/2331186X.2024.2433814>
- [27] Mollick, E., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts*. arXiv Preprint arXiv, 2306, 10052.
- [28] Monika, M., & Suganthan, C. (2024). A study on analyzing the role of ChatGPT in English acquisition among ESL learners during English language classroom. *Language*, 8(2). https://www.bodhijournals.com/pdf/V8N2/Bodhi_V8N2_013.pdf
- [29] Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- [30] Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage Publications.
- [31] Ringo, D. S. (2025). The effect of generative AI use on doctoral students' academic research progress: the moderating role of hedonic gratification. *Cogent Education*, 12(1), 2475268. <https://doi.org/10.1080/2331186X.2025.2475268>
- [32] Roustio, A. (2025). Conversation shaper: GenAI tools can revive and revise writing center discussions of attribution, authorship, & plagiarism. *The Peer Review*. Retrieved July 9, 2025 from <https://thepeerreview-iwca.org/issues/issue-9-1/conversation-shaper-genai-tools-can-revive-and-revise-writing-center-discussions-of-attribution-authorship-plagiarism/>
- [33] Sandstead, M., & Kibler, A. (2025). Voice in L2 writing in the age of AI. *Journal of Second Language Writing*, 69, 101212. <https://doi.org/10.1016/j.jslw.2025.101212>
- [34] Stowe, G. (2023, September 26). *September 2023: Where are we now with AI in writing centers?* WLN Blog. Retrieved November 16, 2025, from <https://wlnconnect.org/2023/09/26/september-2023-where-are-we-now-with-ai-in-writing-centers/>
- [35] Tsufim, F., & Pomerleau, L. (2024). More is less? Using generative AI for idea generation and diversification in early writing processes. In *Teaching and generative AI: Pedagogical possibilities and productive tensions* (Paper No. 37). Utah State University Digital Commons. Retrieved November 16, 2025, from <https://uen.pressbooks.pub/teachingandgenerativeai/chapter/more-is-less-using-generative-ai-for-idea-generation-and-diversification-in-early-writing-processes/>
- [36] Zhang, Z., & Huang, X. (2024). The impact of large-language-model-based chatbots on second-language vocabulary acquisition. *Heliyon*, 10(3), e25370. <https://doi.org/10.1016/j.heliyon.2024.e25370>

Maria Eleftheriou is an Assistant Professor of English at the American University of Sharjah, where she heads the Writing Center and oversees its peer tutor training program. She teaches undergraduate courses in academic writing and research and also teaches in and supervises theses in the MATESOL program at the graduate level. Her scholarship explores writing center pedagogy, multilingual tutoring, online writing instruction, and the role of AI in academic writing. Her publications include work in the TESOL International Journal, Contemporary Educational Technology, the WAC Clearinghouse, Composition Forum, and Palgrave Macmillan. Dr. Maria Eleftheriou's email is meleftheriou@aus.edu and ORCID iD is <https://orcid.org/0000-0001-5453-0590>

Muhammad Ahmer is a Bachelor's of Computer Science Graduate Student from the American University of Sharjah in the United Arab Emirates. He specializes in machine learning, artificial intelligence, and data science research, with a particular focus on applied data science and statistics as well as studying the impact of artificial intelligence in education. In addition, he was a tutor and IT specialist at the university's writing center for 2 years, gaining firsthand experience with artificial intelligence within the writing center. Muhammad Ahmer's email is b00087698@alumni.aus.edu and ORCID iD is <https://orcid.org/0009-0002-6832-5752>

Voices on the Margins: An Intersectional Study of Women Oppression and Agency in Jasmin Darznik's Memoir *The Good Daughter*

Aishwarya Mary Thomas

Department of English, School of Social Sciences and Languages, Vellore Institute of Technology, Vellore, India

Mohamed Sahul Hameed M. A*

Department of English, School of Social Sciences and Languages, Vellore Institute of Technology, Vellore, India

Abstract—This study employs intersectionality theory to analyse Jasmin Darznik's memoir *The Good Daughter: A Memoir of My Mother's Hidden Life* (2012), focusing on the multifaceted oppression of Iranian women shaped by the intertwined forces of patriarchy, religion, culture, and class. This topic is significant as it illuminates the complex sociopolitical structures that marginalise Muslim women in Iranian society, where gender identity and agency are constrained by overlapping systems of power. Women's personal histories are often erased in rigid patriarchal societies. Darznik's memoir serves as a vital narrative that challenges these oppressive frameworks by giving voice to the silenced histories of her mother and her grandmother. Through storytelling as an act of feminist resistance, the memoir confronts the secrecy and shame imposed by societal norms that suppress women's autonomy and their agency. The memoir exposes the multilayered violence endured by women, including forced marriage and honour codes, while highlighting subtle and overt forms of resilience such as education, self-expression, emotional endurance, and intergenerational solidarity. By critically engaging with Darznik's text, this paper highlights the importance of reclaiming marginalised voices and histories to foster an understanding of gendered experiences under intersectional oppression. It challenges the dominant sociopolitical frameworks confining women and affirms their right to autonomy, dignity, and voice. This research contributes to feminist scholarship by illuminating the transformative potential of personal narratives in advancing the understanding of Iranian women's lived realities and challenging dominant discourses.

Index Terms—intersectionality, patriarchy, Iranian women, feminist resistance, gendered oppression

I. INTRODUCTION

Feminism originated from the perception that there is something wrong with society's treatment of women. (Tuttle, 1987). The movement seeks equal political and social rights for women and men. The main common theoretical assumption shared by all branches of the movement is that there has been a historical tradition of male exploitation of women (Robertson, 1993). Porter (1991) adopts a general definition of feminism that sees it as a perspective that seeks to eliminate the subordination, oppression, inequalities, and injustices women suffer because of their sex.

Throughout history, women have endured various forms of oppression, including gender-based violence. The governance of a nation is influenced by its religious beliefs, cultural norms, patriarchal systems, and political frameworks, all of which intersect with factors such as class, gender, and religion. This complex interplay often results in women experiencing unique forms of discrimination, particularly when they belong to multiple marginalised groups, a phenomenon that traditional research often overlooks by isolating individual identities (Chakrabarti, 2023).

Gender identity often becomes a rubric for discrimination, restraining women's roles within the four walls of the house. Except for women with class privilege, most women are left without a choice but to leave education early and marry young which becomes an obstacle to their growth and development. In the modern world, where women advance based on their merit, patriarchy creates obstacles for women to progress in society. Walby defines "patriarchy as a system of social structures and practices in which men dominate, oppress and exploit women" (Walby, 1989, p. 20). Oppression can be understood as the abuse of power that systematically restricts freedom and undermines individual agency.

This subjugation extends across political, economic, social, and cultural domains, reinforcing a societal structure in which men are perceived as inherently superior (Sultana, 2010). This hierarchical organisation often manifests in the imposition of rigid masculinity and femininity character stereotypes, further entrenching unequal power dynamics between genders. Such pervasive systems of control and exploitation significantly impede women's advancement and development by limiting their opportunities and autonomy within the public and private spheres (Sultana, 2010).

When women are subjected to domestication and their intellectual autonomy is constrained through the regulation of sexuality, their rights are correspondingly restricted. In response, women pursue emancipation and resist oppressive

* Corresponding Author. Email: mamohamedsahulhameed@vit.ac.in

practices as culturally embedded forms of agency. Male violence is often condoned and legitimised within the marriage institution. This structural imbalance denies women the freedom to exercise their own agency and make choices, frequently confining them to domestic roles and impeding their pursuit of education or professional development (Sharma & Tyagi, 2025). The intertwining of religious and cultural norms has perpetuated male domination over women.

This oppressive environment shaped the lives of the author's mother and grandmother, both raised in Iran, in comparison to its influence on the author, Jasmin Darznik, who moved to the US at the age of three during the Iranian Revolution. This geopolitical upheaval, coupled with prevailing cultural norms and religious interpretations, profoundly shaped the familial dynamics and individual experiences of women within their community, as it often does in societies undergoing significant societal transformations (Rasheed, 2023).

The term 'female autonomy' refers to the ability of women to make informed decisions about their lives, free from coercion or constraint. By promoting female autonomy, women can exercise their agency and make choices in their best interests. Gender equality is essential for promoting female agency, as it creates a society in which women are valued and respected and can make autonomous choices. Intersectional feminist theory is a critical framework that examines how different forms of oppression, such as racism, sexism, homophobia, and classism, intersect and compound to lead to unique experiences of marginalisation. This paper explores the impact of various forms of oppression on the lives of Iranian women and examines how they assert agency and navigate the constraints imposed by a patriarchal society using Kimberle Crenshaw's intersectional framework. Numerous writers from the diaspora have crafted memoirs that delve into their personal experiences and the silence and trauma passed down through generations, utilising their writing as a tool for narrative resistance against a patriarchal setup.

As Butler (1990) points out in relation to gender identity, it is problematic to discuss women in general because "gender intersects with racial, class, ethnic, sexual and regional modalities of discursively constituted identities. As a result, it becomes impossible to separate out "gender" from the political and cultural intersections in which it is invariably produced and maintained (p. 3). It is crucial to address the subjugation of women within the framework of cultural patriarchy and the intricate identities it is associated with.

Jasmin Darznik is an acclaimed Iranian-born American writer and professor who intricately explores the themes of intersecting gendered and cultural oppression in her debut memoir *The Good Daughter* (2012). Born in Tehran in 1973 and partly raised in the United States, Darznik's life bridges cultures, and her narrative style blends intimate family history with the broader socio-political contexts of Iran before the 1949 revolution. Jasmin Darznik addresses the sociocultural oppression of women, the quest for identity and the struggle for agency, the reclamation of narrative through storytelling, and the small acts of rebellion that assert autonomy within the intersecting frameworks of gender, class, and religion. Hooks argues that speaking from the margin is itself an act of power. Breaking the silence on racism, sexism, and violence is a central act of agency. At a time when there is increasing debate on women's liberation and oppression among many Muslim nations, this memoir invites a nuanced understanding of the traditional, religious, and cultural prohibitions that women face with respect to their gender and how they resist, adapt, or reconstruct their identities through narrative. Female agency appears in constant struggle with "social structures, social institutions, cultural norms and specific groups" (Charrad, 2010) throughout the memoir.

This study also reveals that early Muslim women resisted a religious system that reinforced male domination. In fact, Prophet Muhammed was more emancipated with respect to women than most men of his time. "Islam doesn't make any difference between man and women in the journey from this world towards al-Haqq (the truth towards God)" (p. 172). However, interpretations of Islamic texts and cultural practices have often led to patriarchal systems that subjugate women, necessitating a critical examination of how these traditions are understood and applied (Zubair & Zubair, 2017). Despite these interpretations, many Muslim women actively challenge dominant narratives by reclaiming their agency and engaging in acts of resistance, frequently through creative expression and scholarly work. This active resistance often involves navigating complex intersectional identities, as Muslim women occupy multiple positions influenced by gender, class, ethnicity, and religion, further complicating their bids for selfhood and challenging the established norms (Hussain, 2019). Common forms of injustice and oppression experienced by Muslim women in South Asia include the denial of inheritance rights, the rejection of women's opinions in marriage and divorce, the dissuasion of women from pursuing higher education, and the disregard of women's opinions in social and family matters, all of which are related to agency and self-expression (Ahmed, 2025). Women's entry into the public world is often denied. They need freedom from the bonds of custom and gender stereotypes to be free of societal constraints. In most families, girls are married around thirteen or fourteen before they can complete their education. Iranian diasporic women writers explore these themes in their memoirs and novels.

According to Uchem (2007), gender is a society's expectation of how boys and girls, men and women, in a given culture, ought to feel, look, or behave. Gender inequality occurs through negative attitudes, beliefs, and behaviours that devalue, denigrate, stigmatize, or restrict females in comparison with their male counterpart, resulting in chronic identity trauma among females (Kira et al., 2015). It justifies the objectification of women and the male gaze and women being perceived as instruments of male gratification (Rawat, 2014). In their book *The Self We Live By*, Holstein and Gubrium (2000) argue that identity is not fixed or essential but constructed through narrative practices in everyday life. They emphasize that the "self" is formed through the stories people tell about themselves. These narratives are shaped by social institutions, cultural scripts, and power relations. Understanding the various intersecting factors that influence women is

crucial for comprehending their impact on shaping women's identities. This paper seeks to find a balanced approach to fostering an inclusive society that promotes women's wellbeing by contrasting the conditions in which women thrive and contribute equally to the community with those where women deteriorate and experience mental and physical unhappiness.

II. LITERATURE REVIEW

This review synthesises existing scholarship on women's oppression and agency, particularly within intersectional frameworks, to contextualise Jasmin Darznik's memoir, *The Good Daughter: My Mother's Hidden Life*. Numerous studies have identified the multifaceted nature of women's oppression and agency, often highlighting the complex interplay of social, cultural, and political factors. This body of work frequently examines the marginalisation of women in patriarchal societies and the constraints on their agency and identity.

Research on women's empowerment consistently points to various factors hindering it while identifying successful strategies, such as education, access to resources, and political participation (Reshi & Sudha, 2022). Scholarly findings also show that women's engagement with Western feminist thinking is often complex, with many women wishing to distance themselves from traditional images of subservient women while also harbouring reservations about certain aspects of white feminisms (Zubair & Zubair, 2017).

From an intersectional perspective, research emphasises that social categories such as gender, race/ethnicity/immigrant background, and social class are interwoven, forming a "social location" that shapes lived experience at the individual level (Keller et al., 2023). Intersectionality, a theoretical framework originating from Black feminist theory and activism, named by Kimberlé Crenshaw, investigates how interlocking systems of power and oppression at the societal level influence the lived experiences of historically and socially marginalised groups (Guan et al., 2021; Hanson & Fletcher, 2021). It argues that identities such as gender, race, sexuality, and other markers of difference intersect and reflect larger social structures of oppression and privilege, such as sexism, racism, and heteronormativity (Rice et al., 2019). This framework posits that multiple social categories are not independent but rather interact and mutually shape each other, producing "qualitatively different meanings and experiences". Consequently, intersectionality allows for a complex analysis of people's lived experiences that considers not only various marginalizations but also how these "marginalizations and positionalities intersect in order to create unique situations" (Jorba & López de Sa, 2024).

Feminist theorists argue that feminism, while defying a single fixed definition, generally concerns itself with the identity of women, both philosophically and socially, and with achieving political, legal, and economic equality not only for women but also for all minority and marginalised groups. Feminist theory is understood as a "set of ideas critical of the situation confronting women and offering solutions for improving, if not revolutionising, their situation" (Oyeleye, 2022). It also functions as "a concept that can encompass both an ideology and a movement for sociopolitical change based on a critical analysis of male privilege and women's subordination within any given society" (Oyeleye, 2022).

Abla (2020), in her paper *Diasporic Iranian Women's Life Writing: An Analysis Using a Transnational Feminist Lens*, posits that authors like Darznik reclaim women's agency by giving them a voice through writing. Abla (2020) states that "When women become the narrators of their own stories it becomes an act of resistance towards patriarchal narrative and help in reclaiming their agency in both private and public spheres". Her analysis examines the themes of resistance, marginalisation, and intersectionality in narratives, portraying the struggles of Muslim women in Iran and their experiences of liberation in Western countries. Iranian-American women life writers, despite facing xenophobia and pressure to argue against widespread stereotypes about Iran, portray themselves as both suffering and strong, resisting the demonization of their homeland (Naghbi, 2016). Gondal and Hatta (2024) in *Religion and Gender Studies*, argue that the relationship between religion and gender is complex, shaped by social, historical, and cultural factors, with significant overlap between religious borders, institutions, and various identities like sexuality, race, and gender. Shami, Ijaz, and Akhtar note that women in patriarchal settings face repressive situations and are expected to conform to male-dominated societal norms. They highlight women's desire to "re-think the social structures that deepen their otherness and revitalize male dominance," leading to opposition against prevailing social structures when women pursue agency (Shami et al., 2023). Sara Salem, in "Feminist Critique and Islamic Feminism," offers a critical analysis of Western feminism and its tendency to overlook diverse experiences, advocating for a nuanced understanding where religion and autonomy can coexist through Islamic feminism.

Intersectional research underscores the role of analysing how multiple grounds of identity interact to construct the social world, especially in understanding layered discrimination (Ndobo et al., 2022; Nielsen et al., 2025). Singh (2019) in *Gender and Middle East: An Intersectionality Perspective*, emphasises the need to adapt traditional intersectional methodologies to capture the unique experiences of women in the Middle East, particularly concerning gender roles and societal expectations. Nisa et al.'s (2024) work similarly uses intersectionality to explore how women navigate "layered discrimination at the intersections of various identities, such as gender and religion, within dominant patriarchal power structures".

Memoir or life writing has often been used as a medium to explore and critique, providing a powerful platform for marginalised voices. For instance, Azar Nafisi's *Reading Lolita in Tehran* provides an account of her lived experience as a literature professor in Iran during the Islamic Revolution, using literature as a metaphor for the oppression of women and as a means of escape, resistance, and empowerment. Nafisi recounts personal experiences of suppression by institutional and phallogocentric oppressive elements, including forced veiling and curriculum control (Shami et al., 2023).

This genre allows for the expression of diverse multicultural identities and serves as a medium for writers to share their experiences both in their home country and in the West. As Cixous suggests, "Woman must put herself into the text—as into the world and into history--by her own movement," an encouragement that inspires the writing of memoirs as an "expression of our movement from object to subject—the liberated voice" (Shami et al., 2023). Diasporic/transnational memoir reflects diverse multicultural identities, and these life narratives allow the voices of Middle Eastern women to be heard, challenging their legitimacy (Shahbazi, 2019).

Previous research has examined themes such as exile, cultural alienation, broad intersectional analysis, gender and religion, or feminist critique within specific memoirs or socio-political contexts. Studies have explored how authors like Darznik reclaim women's agency by giving them voice through writing, framing it as an act of resistance to patriarchal narratives. Shahbazi's thesis on Iranian and Iraqi women's life writing adopts a decolonial approach to memoir, viewing it as a site of critical and political inquiry into lived experience and the multiplicity of selves, focusing on multiple transnational memoirs. However, a discernible gap persists in research specifically exploring how women strategically employ agency to pursue educational aspirations, particularly within memoir writing, where the interplay between claiming agency and confronting suppressive social structures remains under-investigated.

This study fills this gap by offering a comprehensive and in-depth examination of how Jasmin Darznik navigates identity, agency, and resistance amid the intersecting factors of gender, class, race, religion, and migration in *The Good Daughter* (2012). While the existing literature has addressed the rethinking of societal structures of dominance, there remains a critical gap in research concerning the explicit, textual relationship between intersectionality, identity, and the act of writing within this memoir as a performative and empowering form of agency and reclamation of voice. This study addresses this key gap by applying an integrated intersectional framework and close reading of *The Good Daughter*, shifting focus from general identity politics to a detailed textual analysis of how the characters' resistance strategies are articulated and how the memoir itself functions as a tool for reclaiming narratives and asserting autonomy. This work broadens the critical conversation around Darznik's memoir using intersectionality theory and positions it within a more inclusive framework of resistance, identity, and belonging.

III. METHODOLOGY

This study employs a qualitative methodology to examine the lived experiences of women in a memoir through the lens of Kimberlé Crenshaw's Intersectionality Theory. This approach offers a detailed investigation into how the experiences of Iranian women and the identity of characters in the memoir are shaped by the complex interactions of gender, cultural background, class, and patriarchy, moving beyond a singular perspective to capture the intricacies of their quest for agency and self-reclamation through the medium of storytelling. This research involves a close reading of the text to uncover various forms of oppression impacting women's identities and how they challenge societal constraints to assert their agency and reshape dominant narratives.

Coined by African American legal scholar Kimberlé Crenshaw in the 1980s, intersectionality challenges mainstream feminism's universalising tendencies by centring the experiences of multiply marginalised individuals. Originating from Critical race studies and Black feminist thought, intersectionality, has garnered significant prominence across diverse academic fields, including women and gender studies, sociology, and race studies (Salem, 2013). Crenshaw's seminal articles, "Demarginalizing the Intersection of Race and Sex" (1989) and "Mapping the Margins" (1991), introduced intersectionality into the academic discourse. A fundamental tenet of intersectionality is that "the intersectional experience is greater than the sum of racism and sexism" (Crenshaw, 1989). This means that multiple intersecting identities, such as gender, race, sexuality, and other markers of difference, produce unique experiences of oppression and privilege that cannot be understood by examining categories in isolation.

As a methodology, intersectionality mandates that research be conducted "from the perspective of the marginalized, as their view of society is more comprehensive" (Salem, 2013). This approach challenges the assumptions of neutrality and objectivity, particularly in legal and social systems, by revealing how systems of power operate through the intersection of various social categories. To effectively implement intersectionality as a methodology, researchers should deconstruct single-axis frameworks and actively replace analytical approaches that isolate categories like race or gender with a framework that demonstrates their dynamic interaction (Salem, 2013).

Intersectionality helps centre marginalized experiences and prioritize the lived experiences of individuals who face compounded oppressions, recognizing that their unique perspectives offer profound insights into systemic inequalities. This includes understanding how women of colour "could see, feel, and experience how the treatment of their bodies as simultaneously raced and gendered shaped the contours of their subordination" (Kelly et al., 2021). Intersectionality as a theory helps to analyse power dynamics and critically examine how intersecting identities reflect and reproduce larger social structures of oppression and privilege, ensuring the inclusivity of varying experiences, realities, and identities.

For my qualitative study of Jasmin Darznik's memoir, *The Good Daughter* (2012), employing an intersectional methodology allows a nuanced exploration and multidimensional perspective of how Iranian women's experiences are shaped by the intricate interplay of gender, cultural heritage, class, and diasporic identity, moving beyond a singular focus to capture the layered complexity of women's struggles for agency and self-reclamation, demonstrating how storytelling functions as a means to challenge dominant narratives and assert marginalised voices.

IV. DISCUSSION

A. *Layers of Oppression in The Good Daughter*

The label or trope of the 'Good Daughter' signifies an individual who does not challenge societal oppression and quietly conforms to the norms established by culture and authority, lacking independent thought. It involves adhering to societal expectations without questioning or resisting them.

The various facets of women's identities intersect to shape their lived experiences, primarily through race, class, religion, and tradition. A significant challenge confronting contemporary feminism is the need to address the exclusion of the experiences of women from marginalised groups within the broader female community. This memoir highlights the mistreatment of women, particularly those belonging to the middle class or those without a male head to provide them with support. Women who lack financial resources are often subjected to ridicule and physical abuse by their husbands. Intersectionality refers to the notion that social relations are characterised by multiple overlapping layers of discrimination.

Patriarchy is a deeply entrenched system that regulates power within society. Families seem to be pushing their daughters into firepits through early marriage and trading their lives for money, as Kobra's brother traded her to his opponent in gambling because he had no money to repay. Women count on the love of their husbands as early marriage limits their access to the world, while men take a new wife when they get bored with one. In Sohrab's case, his father abandoned his mother and subsequently entered marriages with other wives, whom he later abandoned to marry other women. This practice is not regarded as prostitution within their cultural framework. However, if women engaged in multiple marriages, it was deemed adultery. In this context, women often resort to witchcraft and consult fortune tellers to maintain their husbands' affection. These women experience profound insecurity as they rely heavily on their husbands and feel compelled to employ various strategies and manipulations to prevent their husbands from engaging in infidelity. Ultimately, both Kobra and Lili attain autonomy and overcome their co-dependency upon realizing that men are unlikely to alter their inherent nature. Kobra reconciles with her marital situation until her husband's demise, whereas her daughter, Lili, insists on pursuing a divorce to reclaim her agency before succumbing to insanity or potential harm. The strategies employed to exert excessive control involved curtailing women's education at an early age, thereby rendering them powerless and reliant on their husbands, and marrying them off before they matured and discovered their identities and desires. Post-marriage, women were isolated from all except their husbands and his family, and any independent outings required their husbands' permission. This isolation from the external world was intended to prevent them from escaping the control and achieve autonomy. Women were subjected to continuous surveillance and monitoring by their family members and faced punishment for non-compliance with established standards, a practice intended to deter resistance and maintain familial dominance over women.

Upon Lili's relocation to Germany for her studies and the commencement of a new chapter in her life, she attains a renewed sense of independence. However, she also encounters instances of racial discrimination, including derogatory remarks and hostile stares while navigating public spaces. Although migration has facilitated her aspiration to be addressed as Madame doctor and to become a respected professional woman with an international qualification, she simultaneously faces discriminatory challenges in foreign land. For Lili, migration to US also does not offer complete emancipation. Though she gains more legal and social freedoms in the U.S., her internalized patriarchy and cultural trauma continue to affect how she raises Jasmin and there also she faces some form of racism while she speaks in English as part of the Iranian diaspora. Some customers come to stay in her hotel and doesn't pay and cause trouble.

Patriarchy is normalised in domestic spheres under the guise of tradition and honour. In the memoir, women in Sohrab's house rarely left the house unless it was for a family event or pilgrimage to preserve family honour. "The large colony of sisters and aunts and mothers and grandmothers within never left except to attend a wedding or funeral close by or else to make a pilgrimage to a martyr's shrine. And for that they always travelled with their men" (Darznik, 2012, p. 9). Women travelling alone and unaccompanied by men were seen as vulnerable targets.

One evening, Kobra's brother, Ali Ahmed, who was known for his gambling habits, made a proposal that would change Darznik's grandmother's life forever. After suffering his largest financial loss in what would become a notorious gambling career, Ali Ahmed turned to his opponent and offered, "You can marry my sister." He didn't mention her name, only referring to her as "the youngest one" (Darznik, 2012, p. 13). Darznik's grandmother Kobra was sold by her brother to his friend and opponent to settle his gambling debts. He did not even consider her opinion or ask for it. He announced her marriage and informed the family.

Darznik writes about married women's predicament under the authority and control of their husbands. Most evenings, Sohrab joined his friends at gatherings where the women were very slim, dressed in Western attire, and far more attractive than Kobra. When she expressed a desire to accompany him, he remarked that she appeared aged and untidy in her chador, and he would feel ashamed to bring her along. She even suggested going without her veil or just a simple headscarf, despite this being almost as uncomfortable for her as walking the streets unclothed. Nevertheless, he continued to go out by himself (Darznik, 2012, p. 18).

Women from elite social strata were subjected to differential treatment. Sohrab is associated with women who are modern, educated, and sophisticated, yet he feels ashamed to introduce his wife at gatherings of the affluent. Despite Kobra's youth, her husband neglects her because of her modest background. Even when she compromises her principles to please him, she is met with rejection. She was relegated to the traditional role of a domestic servant, attending to his daily needs and assisting his mother and aunts with the household chores. He did not provide her with the financial means

to groom herself or purchase new clothing, yet he criticised her as unkempt and shabby. Furthermore, he treated her with disrespect, resorted to physical violence, and confined her to the house while parading himself like a peacock outside. She is impregnated at fourteen, and before she recovers from childbirth, she is already pregnant for the second term. This shows her husband's double standards and lack of concern for his wife. Whether it is Kobra's brother or her husband, their attitude towards her reflects the lack of agency experienced by women in their culture.

"During that time, a young bride was not referred to by her own name. She was simply called the aroos, or 'the bride,' and she only adopted her husband's name after her mother-in-law passed away" (Darznik, 2012, p. 14). The entire system was unjust. Once a woman married, she had no status beyond being referred to as the bride until her mother-in-law passed away, at which point she could take her husband's name. The invisibility of women is common in their culture. In *The Good Daughter*, Darznik's narrative reveals how the institution of marriage operates as a site where male authority is normalised and, at times, shields forms of male violence through cultural expectations of obedience, endurance, and familial honour. Such legitimisation constrains women's agency, compelling silence and compliance while rendering resistance socially transgressive.

B. Religion, State and Gender: Interlocking Systems of Control

Modern feminist writer Nawal El Sa'adawi said the political and ideological use of religion is wrong and she sees writing as a weapon of carrying this message across because writing to her "draws its authority from the autocratic power exercised by the ruler of the state, and that of the father or the husband in the family ... [and it thus] became an act of rebellion against injustice exercised in the name of religion, or morals, or love (Shukri & Owoyemi, 2016, p. 11).

Sohrab's mother was a generous and astute woman. Despite her husband being a womanizer who married three wives and then left her without any financial support, she managed the entire household on her own and raised Sohrab and his sisters. By the time Sohrab turned two, his father had already left their home on Avenue Moniriyeh and taken three more wives. When these three women, who had also been abandoned by her husband, showed up at her door one by one, she graciously took them in. They lived together like sisters and worked alongside her (Darznik, 2012, p. 14). For Sohrab too, his first marriage was merely an agreement. He later lived with Simin, a twice-divorced barren woman, as his second wife, although he continued to send money home.

As the Muslim world transitioned into the modern era, the struggle for the rights of Muslim women took a new direction. Following the example set by the modern Western world, Muslim women began to unite, forming associations and movements to demand their rights from society, particularly from their spouses, fathers, brothers, and uncles. Women are not criminals that should be confined in women's quarters. In her book "Veiling and Unveiling," Nazira Zayn al-Din examines the comparative progress of veiled women in Muslim nations and their unveiled counterparts in more advanced nations. She argues that unveiled women are more progressive, educated, and responsible compared to veiled women, who are often illiterate and non-progressive, confined and segregated within society. Al-Din contends that the practice of veiling was historically imposed by Muslim men who failed to adhere to the Qur'anic injunctions regarding modesty, which apply to both men and women. Due to the patriarchal nature of Muslim society, women were disproportionately burdened with the requirement to wear the veil, which she interprets as a symbol of oppression. Furthermore, al-Din asserts that women are more qualified than men to interpret Qur'anic verses concerning women's affairs, as these verses directly address women. Islamic verses are characterized by their fluidity and openness to interpretation, as many of them are inherently ambiguous. This ambiguity allows for diverse interpretations to be applied across various contexts, cultures, times, and situations (Hamdeh, 2020).

The way they treat women tells a lot about society. When a society cancels women, it becomes a barbaric society. If a man commits a sin against women, it's still considered women's fault, and they push women to the periphery without any choice. Sohrab ended his marriage to Kobra impulsively, bypassing any official process. During the time of Sohrab and Kobra's union, a woman who resided away from her husband's home could be reclaimed within three and a half months, effectively cancelling the divorce. This rule was ostensibly to confirm she wasn't leaving while pregnant with his child—an act considered akin to theft—but in Kobra's situation, the rule would serve a different function (Darznik, 2012, p. 26). Women were always asked to silently accommodate no matter what her husband did. Several times, Kobra would be cast out by her husband after a fight, and her mother-in-law and Sohrab's sisters would pick her up from her mother Pargol's house before the three-month period was over so that she could come back and take care of the children. In their culture, only the father held the legal parental right to custody of the children.

Raday (2016) posits that Islamic cultural practices, such as the requirement of modesty through the hijab, can be perceived as mechanisms for exerting patriarchal control over women's behaviour and bodies (p. 306). Similarly, the conduct and appearance of women are shaped by religious traditions that adhere to ideals of modesty and preserving family honour, which frequently mirror entrenched patriarchal power structures within religious institutions. In Islamic countries such as Saudi Arabia, Afghanistan, and Iran, women are required to wear veils upon reaching puberty. In Iranian culture, women without veils have historically been stigmatised as immoral. Wearing headscarves has been culturally mandated from the age of ten. However, during the reign of Reza Shah, reforms were introduced, including the promotion of women's education, the prohibition of the veil, and the adjustment of the legal minimum age for marriage from nine to sixteen. Despite these reforms, families frequently circumvented the law or restricted girls from leaving their homes. Lili, for instance, was married at the age of thirteen to avoid spending another year unveiled in public. Her aunts and grandmother, who were deeply religious, chose not to leave their homes when the veil was outlawed in 1936 (p. 43).

C. *Silenced Voices and Reclamation of Agency*

"A divorcee, it was said, surrendered to seduction as easily as a ripe peach slid down the throat. Sohrab predicted that any young man who discovered Lili's status would make advances on her- or much worse- and since she was no longer a virgin, there would be no recourse for such actions. The only way to protect Lili now would be to keep her at home" (p. 114). According to Iranian standards, divorced women are considered damaged goods. Anyone could eye her, make sexual advances towards her, and exploit her on the street since she was no longer a virgin or the property of her husband. Nobody would come to her defense if anything evil occurred to her. As a result, her father had to either keep her at home, send her to a place where her past was unknown for education, or send her abroad. "The rest of the plan went like this: If he could not restore her to the status of a respectable women, then Sohrab was certain that a high school diploma, followed by some kind of occupational training, would shield her from the curses and insults that would trail a divorcee through the rest of her life" (p. 118). Being a divorcee implied a lifetime of neglect and disrespect; to protect his daughter's reputation after numerous pleas from her male cousin, Lili's father Sohrab decides to further her education and transform her into a professional woman. Eventually, he sends her abroad to earn a foreign degree. This decision alters Lili's destiny, unlike the other women in her family. Upon her return to Iran following her father's death, as she begins her job search, her aunts seek another marriage proposal for her and urge her to remain at home. In her country, only women who had no other options worked. Her aunts viewed it as akin to selling oneself on the streets. In their eyes, a divorced woman can only be redeemed through marriage.

"When Lili returned from Europe to become a bride in the early sixties, she did so by force of her own considerable ambitions, brandishing a foreign diploma that compelled everyone to call her Madame Doctor- the first of the family and the first for many years yet. To the even greater astonishment of her cousins, aunts, and stepmothers, she also brought with her a handsome, well-mannered, blue-eyed farangi eager to live in the country and also to convert to Islam in order to marry her" (Darznik, 2012, p. 197). Lili takes charge of her life and goes back to Germany, resumes her education, becomes a professional midwife, and gets engaged to a German engineer, bringing him back to the country. When one woman takes control of her life, she empowers other women to do so. One person's stand against injustice can eventually move an entire nation. Women who follow the rules and marry into families at a young age are treated with contempt after their divorce. In 1962, it was a very bold move for a woman from native Iran to marry a foreigner, but she successfully crosses all the ordeals.

"Why do you insist on wearing white?" her aunts and stepmothers had asked her with varying degrees of approbation. "It isn't done for a second marriage; you will only call attention to yourself by wearing a white dress." As these were the very relatives who'd hastened her marriage to Kazem, Lili suppressed the urge to scream, curse and strike. She did, however, go so far as to remind them that her first wedding dress, the one Kazem's relatives had chosen for her, had not even been white but pale blue. Besides, Lili told them, the marriage was Johann's first, and on his account, she would consider nothing but white" (Darznik, 2012, p. 204).

Despite her family's attempts to dissuade her from wearing white for her second wedding, she remained resolute. This is her first marriage of her choice, and the previous one nearly cost her life.

D. *Writing as Mode of Resistance*

Bell Hooks often says speaking from the margin is itself an act of power. Breaking silence about racism, sexism, and violence is central towards agency. Through writing the stories of three women in their generation, she regains their voice and reclaims the agency. Jasmin was unaware of her mother Lili's first marriage and her half-sister Sara living in Iran until she stumbled upon a photograph of Lili's wedding to Kazem after her German father passed away. After a period of silence, Lili chose to share the story of her grandmother, her mother, and herself, highlighting how these three women, in their unique ways, challenged stereotypes and refused to succumb to their destinies. They took charge of their lives by resisting oppression, despite cultural pressures to silence them, limit their freedom, and dictate their choices. Documenting their stories served as a means to recognize their lives and voices, empowering them through the written word. Darznik reconstructs her mother Lili's hidden life, revealing a past that had been buried under layers of cultural shame and family secrecy. Lili's story is shaped by gendered oppression within Iranian patriarchy, where her girlhood was regulated through honor, obedience, and early marriage. Yet she fights back and refuses to blend in and get the life she wants. By writing this memoir, Darznik reclaims her narrative as well along with histories of other women in her family. The memoir shows how diasporic women are not automatically liberated by migration but instead must wrestle with inherited burdens and culturally embedded silence.

For Jasmin, understanding her mother's past becomes an act of self-discovery. She reclaims her mother's silenced narrative enabling a dialogic healing between past and present. Both Lili and her mother Kobra prove through the latter half of their lives that given freedom and the right training, these women can be fiercely independent yet run their homes well even if their husbands lack the responsibility of running the household or dies. Its Lili who does odd jobs and sends money to her brother abroad for his studies when her father dies and Kobra who also does what she knows best to sustain her household. These women more autonomous and freer only after the death of the only patriarchal figure in their family. Storytelling here emerges as an act of feminine resistance to societal constraints and asserting the autonomy of self.

V. CONCLUSION

This paper elaborates how patriarchy is entrenched in societies and how women's identity is shaped by intersection of multiple factors such as gender, class, tradition, and religion. Feminist writing has been crucial in opposing institutional injustice and supporting social change. Feminist writers have given voice to women who don't have much by telling critical stories that show how legal, economic, or educational barriers affect women. This has led to resistance and empowerment at the same time (Sharama, 2020, p. 103). Through writing the marginalised voices reclaim their agency. Jasmin Darznik advocates for a social change in this context, through a change in the treatment of women through education, empowerment, support systems, political and religious reforms and assertion of female identity as equal in status to men. Women should have the freedom to pursue what they want without restraints and feelings of shame towards making their own choices. Through her memoir, she demonstrates how storytelling functions as an ultimate act of liberation in challenging dominant narratives and asserting silenced and marginalized voices to reclaim their agency. The study has shown that oppression operates through a combination of multiple factors rather than a single problem such as domestic, religious and socio-legal forces. The identity is a fragmentation of cultural construct and acts of resistance. It proves that agency can be reclaimed through literacy, migration, and storytelling. Future research could explore comparative analyses of feminist resistance across different religious and cultural contexts, investigate the impact of digital media on the reclamation of agency among marginalized women, and further examine the role of intersectionality in shaping contemporary feminist movements. Such studies would deepen insights into the evolving dynamics of gender, identity, and power in diverse socio-political landscapes.

REFERENCES

- [1] Abla, F. (2020). *Diasporic Iranian Women Life Writing: An Analysis Using a Transnational Feminist Lens* [Doctoral dissertation]. Concordia University.
- [2] Ahmed, S. (2025). *Exercising Transformative Agency: South Asian Muslim Women Living in the United States* [Doctoral dissertation]. The Chicago School of Professional Psychology.
- [3] Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. New York, NY: Routledge.
- [4] Chakrabarti, M. (2023). *The Predictiveness of Gendered Racism on the Well-Being of Women of Color* [Doctoral dissertation]. The University of North Dakota.
- [5] Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory, and antiracist politics. *University of Chicago Legal Forum*, 1989(1), Article 8, 139–167. <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>
- [6] Darznik, J. (2012). *The Good Daughter: A Memoir of My Mother's Hidden Life*. Random House.
- [7] Essien, S. P. (2020). Feminism in Culture and Religion. *World Journal of Innovative Research*, 8(5), 30-36.
- [8] Gondal, A. Q., & Hatta, Z. (2024). Religion and Gender Studies: Examining the Intersection. *Journal of Development and Social Sciences*, 5(1), 302-310.
- [9] Guan, A., Thomas, M. D., Vittinghoff, E., Bowleg, L., Mangurian, C., & Wesson, P. (2021). An investigation of quantitative methods for assessing intersectionality in health research: A systematic review. *SSM-population health*, 16, 100977. Elsevier BV. <https://doi.org/10.1016/j.ssmph.2021.100977>
- [10] Hanson, C. L., & Fletcher, A. J. (2021). Beyond the Trinity of Gender, Race and Class. *European Journal for Research on the Education and Learning of Adults*, 12(2), 135. <https://doi.org/10.3384/rela.2000-7426.3360>
- [11] Holstein, J. A., & Gubrium, J. F. (2000). *The self we live by: Narrative identity in a postmodern world*. New York: Oxford University Press.
- [12] Hooks, B. (2015). *Talking Back Thinking Feminist, Thinking Black*. New York: Routledge Taylor and Francis Group.
- [13] Hussain, S. (2019). Social justice, politics of authorisation and agency: A hybrid theoretical framework to study contemporary Muslim femininity. In *Women's Studies International Forum* (Vol. 74, pp. 188-195). Pergamon. <https://doi.org/10.1016/j.wsif.2019.04.005>
- [14] Jorba, M., & López de Sa, D. (2024). Intersectionality as emergence. *Philosophical Studies*, 181, 1455-1475. <https://doi.org/10.1007/s11098-024-02155-1>
- [15] Keller, L., Lüdtke, O., Preckel, F., & Brunner, M. (2023). Educational Inequalities at the Intersection of Multiple Social Categories: An Introduction and Systematic Review of the Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA) Approach. *Educational Psychology Review*, 35(1). <https://doi.org/10.1007/s10648-023-09733-5>
- [16] Kelly, C., Kasperavicius, D., Duncan, D., Etherington, C., Giangregorio, L., Pesseau, J., Sibley, K. M., & Straus, S. (2021). 'Doing' or 'using' intersectionality? Opportunities and challenges in incorporating intersectionality into knowledge translation theory and practice. *International Journal for Equity in Health*, 20(1), Article 187. <https://doi.org/10.1186/s12939-021-01509-z>
- [17] Kira, I. A., Omidy, A. Z., Fawzi, M. H., Rice, K. G., Fawzi, M. M., Lewandowski, L., & Bujold-Bugeaud, M. (2015). Are the Negative Mental Health Effects of Gender Discrimination (GD) Salient across Cultures? Does Self-Esteem Mediate These Effects: GD as a Continuous Traumatic Stress and the Pathways to Its Negative Dynamics? *Psychology*, 6(1), 93-116. <https://doi.org/10.4236/psych.2015.61009>
- [18] Naghibi, N. (2016). *Women write Iran: Nostalgia and human rights from the diaspora*. University of Minnesota Press.
- [19] Ndobo, A., Fauré, A., Galharret, J., Sarda, E., & Debont, L. (2022). The coping strategies of individuals in multiple jeopardy settings: the case of unemployed older women. *The Journal of Social Psychology*, 164(5), 640-665. <https://doi.org/10.1080/00224545.2022.2143313>
- [20] Nisa, I. S., Shapiulayevna, A. P., & Fikri, M. N. (2024). The role of women in Islamic politics: From the early period to the contemporary era through the lens of Kimberlé Crenshaw's intersectionality perspective. *International Journal of Multicultural Studies*, 5(2), 271–281. <https://doi.org/10.30984/kijms.v5i2.1252>

- [21] Oyeleye, O. A. (2022). Feminist Postproverbial as a Panacea for Decolonising African Feminist Scholarship. *Journal of Higher Education in Africa*, 20(1), 85-106. <https://doi.org/10.57054/jhea.v20i1.2188>
- [22] Porter, E. J. (1991). *Women and moral identity*. Oxford University Press.
- [23] Rawat, P. S. (2014). Patriarchal beliefs, women's empowerment, and general well-being. *Vikalpa*, 39(2), 43-56.
- [24] Reshi, I. A., & Sudha, D. T. (2022). Women Empowerment: A Literature Review [Review of *Women Empowerment: A Literature Review*]. *International Journal of Economic Business Accounting Agriculture Management and Sharia Administration (IJEBAAS)*, 2(6), 1353–1359. <https://doi.org/10.54443/ijeabas.v2i6.753>
- [25] Rice, C., Harrison, E., & Friedman, M. (2019). Doing justice to intersectionality in research. *Cultural Studies Critical Methodologies*, 19(6), 409-420. <https://doi.org/10.1177/1532708619829779>
- [26] Robertson, D. (1993). *The Penguin dictionary of politics* (2nd ed.). Penguin Books.
- [27] Salem, S. (2013). Feminist critique and Islamic feminism: The question of intersectionality. *The Postcolonialist*, 1(1), 1-8.
- [28] Shahbazi, S. (2018). *Decolonial Feminist Epistemologies of Border crossing: A Comparative Study of Transnational Iranian and Iraqi Women's Life Writing*. Retrieved Oct 9, 2025, from https://ses.library.usyd.edu.au/bitstream/2123/20411/1/Shahbazi_sh_thesis.pdf
- [29] Shami, Q., Ijaz, M., & Akhtar, Z. (2023). Attaining Women's Agency: A Study of Women's Repression in Women's Memoir Writing. *Journal of Social Sciences and Humanities*, 31(2), 84–96.
- [30] Sharama, N. (2020). Feminist Literature as a Catalyst for Social Change: Examining Institutional Oppression and Maternal Resilience in English Literature. *International Journal of English and Studies*, 2(7), 99–105. <https://www.ijoes.in>
- [31] Sharma, B., & Tyagi, S. R. (2025). A Study of Women Oppression and Injustice in Indian Theatre. *International Research Journal of Multidisciplinary Scope*, 6(1), 315–323. <https://doi.org/10.47857/irjms.2025.v06i01.02501>
- [32] Shukri, A. M., & Owoyemi, M. Y. (2016). The origin, development and diversity of Islamic feminism: A historical review. *MALIM: Jurnal Pengajian Umum Asia Tenggara*, 17, 1–18. <https://doi.org/10.17576/malim-2016-1701-01>
- [33] Singh, R. (2019). Gender and Middle-East: An intersectionality perspective. *English Linguistics Research*, 8(2), 10–16. <https://doi.org/10.5430/elr.v8n2p10>
- [34] Sultana, A. (2010). Patriarchy and women's subordination: a theoretical analysis. *Arts Faculty Journal*, 36, 1-18. <https://doi.org/10.3329/afj.v4i0.12929>
- [35] Tuttle, L. (1987). *Encyclopedia of feminism*. Pantheon Books.
- [36] Walby, S. (1989). Theorising Patriarchy. *Sociology*, 23(2), 213–234. <https://doi.org/10.1177/0038038589023002004>
- [37] Zubair, S., & Zubair, M. (2017). Situating Islamic feminism(s): Lived religion, negotiation of identity and assertion of third space by Muslim women in Pakistan. *Womens Studies International Forum*, 63, 17–26. <https://doi.org/10.1016/j.wsif.2017.06.002>

Aishwarya Mary Thomas is a full-time research scholar in the Department of English, School of Social Science and Languages, Vellore Institute of Technology, Vellore, India. She completed her undergraduate from Bishop Moore College, Mavelikara and her postgraduate from Pondicherry Central University, Puducherry.

She is currently pursuing her doctoral research in English literature. Her area of interest is Diasporic Literature and Gender Studies. She likes to write articles on humanitarian issues and women centric problems. Her research area focuses on women of colour.

M. A. Mohamed Sahul Hameed is currently an English professor at the School of Social Sciences and Languages, VIT University, Vellore, India. He completed his PhD at VIT, Vellore, India. His main area of specialization are English Language and Teaching (ELT), Literature and Indian Writing in English.

He has published numerous papers in Scopus, Non-Scopus and as book chapters. He has also conducted several training programs and delivered lectures on diverse topics. He is also an author of several books, and his reviews and articles are published in several newspapers and magazines. He can be reached at mamohamedsahulhameed@vit.ac.in; his ORCID iD is <https://orcid.org/0000-0002-8911-6610>.

Call for Papers and Special Issue Proposals

Aims and Scope

Journal of Language Teaching and Research (JLTR) is a scholarly peer-reviewed international scientific journal published bimonthly, focusing on theories, methods, and materials in language teaching, study and research. It provides a high profile, leading edge forum for academics, professionals, consultants, educators, practitioners and students in the field to contribute and disseminate innovative new work on language teaching and research.

JLTR invites original, previously unpublished, research and survey articles, plus research-in-progress reports and short research notes, on both practical and theoretical aspects of language teaching, learning, and research. These areas include, but are not limited to, the following topics:

- Language teaching methodologies
- Pedagogical techniques
- Teaching and curricular practices
- Curriculum development and teaching methods
- Programme, syllabus, and materials design
- Second and foreign language teaching and learning
- Classroom-centered research
- Literacy
- Language education
- Teacher education and professional development
- Teacher training
- Cross-cultural studies
- Child, second, and foreign language acquisition
- Bilingual and multilingual education
- Translation
- Teaching of specific skills
- Language teaching for specific purposes
- New technologies in language teaching
- Testing and evaluation
- Language representation
- Language planning
- Literature, language, and linguistics
- Applied linguistics
- Phonetics, phonology, and morphology
- Syntax and semantics
- Sociolinguistics, psycholinguistics, and neurolinguistics
- Discourse analysis
- Stylistics
- Language and culture, cognition, and pragmatics
- Language teaching and psychology, anthropology, sociology
- Theories and practice in related fields

Special Issue Guidelines

Special issues feature specifically aimed and targeted topics of interest contributed by authors responding to a particular Call for Papers or by invitation, edited by guest editor(s). We encourage you to submit proposals for creating special issues in areas that are of interest to the Journal. Preference will be given to proposals that cover some unique aspect of the technology and ones that include subjects that are timely and useful to the readers of the Journal. A Special Issue is typically made of 15 to 30 papers, with each paper 8 to 12 pages of length.

A special issue can also be proposed for selected top papers of a conference/workshop. In this case, the special issue is usually released in association with the committee members of the conference/workshop like general chairs and/or program chairs who are appointed as the Guest Editors of the Special Issue.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- Description of the topic area to be focused upon and justification
- Review process for the selection and rejection of papers
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- List of potential reviewers if available
- Potential authors to the issue if available
- Estimated number of papers to accept to the special issue
- Tentative time-table for the call for papers and reviews, including
 - o Submission of extended version
 - o Notification of acceptance
 - o Final submission due
 - o Time to deliver final package to the publisher

If the proposal is for selected papers of a conference/workshop, the following information should be included as part of the proposal as well:

- The name of the conference/workshop, and the URL of the event.
- A brief description of the technical issues that the conference/workshop addresses, highlighting the relevance for the journal.
- A brief description of the event, including: number of submitted and accepted papers, and number of attendees. If these numbers are not yet available, please refer to previous events. First time conference/workshops, please report the estimated figures.
- Publisher and indexing of the conference proceedings.

If a proposal is accepted, the guest editor will be responsible for:

- Preparing the “Call for Papers” to be included on the Journal’s Web site.
- Distribution of the Call for Papers broadly to various mailing lists and sites.
- Getting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Author Guide.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

More information is available on the web site at <http://www.academypublication.com/jltr/>

