

An Investigation of the Impact of Task Complexity on ESL Learners' Spoken Language Performance

Amir A. Hassanein

Faculty of Education, The British University in Dubai, Dubai, United Arab Emirates

Emad A. S. Abu-Ayyash

Faculty of Education, The British University in Dubai, Dubai, United Arab Emirates

Abstract—This study investigates the multiple effects of task complexity on language performance and production in English as a Second Language (ESL) classrooms. Through its qualitative design, the present study aims to gain well-founded insights about the potential influence the degree of task complexity has on actual oral productions by ESL learners. In particular, this study construed how the increasing task complexity affected students' oral production. Given that 'oral production' is a notion that involves an intricate web of variables, the present study sought to come to grips with three major components of oral production, which are fluency, vocabulary and grammar. The researchers used the IELTS speaking band descriptor to benchmark the learners' productions. In essence, the study sought to answer three research questions: 1) What is the impact of task complexity on the quality of high achievers' and mid achievers' speech production?, 2) What is the relationship, if any, between learners' language proficiency and their performance in monologic and dialogic tasks?, and 3) Does increasing the cognitive complexity of the task impact oral language performance? The findings revealed that high achievers' grammar, lexical resource and fluency aligned variably with the task complexity, whereas the three elements dropped markedly among mid-achievers.

Index Terms—task complexity, oral production, language accuracy, fluency, language production, ESL classrooms, IELTS speaking

I. INTRODUCTION

This part provides an overview of the present paper, hence dwelling upon the notion of task complexity and its relationship to the production skills, particularly speaking. This part also identifies the research questions that the present investigation sought to answer. Also outlined in this part are the significance and the scope of the present paper.

A. An Overview

Focus on speaking as a language skill to be learned and taught in classrooms has been the cry in an ample body of second language acquisition research. This inclination has, quite understandably, led to the rise of new teaching and learning methods within pedagogical approaches, such as Communicative Language Teaching (CLT) and Task-Based Language Teaching (TBLT). For the latter, a host of scholars (e.g., Skehan, 2003; Van den Branden, Bygate, & Norris, 2009) believe that syllabus designers and teachers involved should understand the nature of the tasks that students will perform in order to plan a lesson with effective instructions to support learning and second language acquisition (SLA). Task analysis is probably the most important part of instructional design for L2 learners. It is rightly believed that TBLT achieves important pedagogical objectives like helping learners to negotiate and comprehend the meaning of language inputs through task instructions or other classmates' interaction with the same task (Baralt & Gomez, 2017) and to negotiate the form (Batstone, 2016). Also, tasks provide a great opportunity for positive corrective feedback from teachers or task partners. ESL teachers gain a reflective benefit by observing the difference of the learner's language input and his/her output; that can help the teacher as a means of formative assessment to help stress the areas of weakness that need more reinforcement.

Theoretically, the cognition hypothesis explains the extent to which task complexity affects the language performance, interaction, and learning that underpin it. The main pedagogic objective of the cognition hypothesis is to know how to design and arrange tasks gradually in order to promote students' learning and language performance. Specifically, the cognition hypothesis, states that there are likely to be positive effects on speech production when tasks are made complex; learners will consequently have an increased accuracy and complexity in order to cope with the complex level of the task. The Cognition Hypothesis agrees with Charles Reigeluth's (1999) 'Elaboration Theory' of how to have a 'holistic approach' to identifying and 'sequencing' the units of instruction. Reigeluth says "The paradigm shift from teacher-centered and content-centered instruction to learner-centered instruction is creating new needs for ways to sequence instruction (p. 427).

At the other end of the scale, Skehan's (1998) 'Limited Capacity Hypothesis' claims that more cognitively complex tasks "consume more attentional resources...with the result that less attention is available for focus on form" (p. 97), therefore, language output like grammar rules, spelling, and punctuation can be less accurate. Due to scarcity of attentional resources, tasks can lead either to increased complexity, or accuracy of production, but not to both. Skehan (1998) therefore recommends that tasks should be sequenced by choosing those with characteristics that lead to fluency, accuracy and complexity, at an appropriate level of task difficulty, as determined by three factors: (1) code complexity, which is described in 'fairly traditional ways', as in descriptions of structural syllabuses, or developmental sequences (p. 99); (2) cognitive complexity, which is the result of the familiarity of the task, topic or genre, and the processing requirements; information type, clarity and organization, and amount of computation required; and (3) communicative stress, which involves six characteristics including time pressure, number of participants, and opportunities to control interaction. These characteristics, Skehan (2002) argues, can be manipulated during task design, and by teachers using tasks "to orient learners away from simply focusing on meaning, but also push them to extending and at the same time achieving greater control over the language" (p. 293).

Looking at the broader context, it does not seem much to claim that public speaking skills and debating skills are gaining an increasing momentum in English language classrooms owing to their importance for learners whose journey outside the educational system lends itself in no small part to a good level of speaking proficiency. Examples of life endeavors that call for appropriate speaking skills include job and immigration applications.

The main focus of the present paper, which is the relationship between task complexity and the speaking skill, was not absent from previous research (e.g. Hyun & Lee, 2018; Vasylets, Gilabert & Manchon, 2017). However, the findings of the previous studies on task complexity were noticeably incompatible. Among the factors responsible for this incompatibility, primacy was given to language proficiency, age, anxiety, the task's cognitive complexity and task familiarity (Cho, 2018; Gilabert, 2005; Rahimpour, 1997; Robinson, 1995; Skehan, 1998). More discussion on this will be presented in the literature review section of this paper.

B. Purpose of the Study

This study generally purports to explore the impact of task complexity on the language production and performance. To do so, this study is set to answer the following research questions:

- What is the impact of task complexity on the quality of high achievers' and mid-achievers' speech productions?
- What is the relationship, if any, between learners' proficiency and their performance on both monologic and dialogic tasks?
- Does increasing the cognitive complexity of the task impact oral language performance?

C. Significance of the Study

Although copious studies have investigated the notion of task complexity in teaching and learning environments, the present study builds on and contributes to the literature in a number of ways. This study examines the relationship between task complexity and the oral production of learners across two academic levels, namely mid-achievers and high-achievers, a tendency that has so far been present at a scanty level in the literature. Similarly, the dearth in the studies that explore dialogical tasks, a major component of the present paper, is also evident in the literature.

D. Scope of the Study

This exploratory study was conducted in a high school that follows the American curriculum in Dubai, United Arab Emirates. on four female participants; the participants were four grade 12 female students. Two of the participants were high achievers, while the other two were mid-achievers; the participants were selected based on their school grades, IELTS scores, and teacher's evaluation and feedback. This would facilitate gaining more insight about the impact of task complexity on two different levels of students.

II. LITERATURE REVIEW

This section reviews the literature on Task Based Language Teaching (TBLT) as it accounts for the theoretical framework of the present investigation. In addition, this section will cover Interactional Competence which is related to the nature of task 3, the mini-debate. The literature will also introduce the general findings of studies about the effects of task complexity on language production or performance by students.

A. Task-based Language Teaching

Bygate et al (2001) define task as an activity which requires learners to use language, with emphasis on meaning, to achieve an objective. In essence, most researchers and linguists agree that the task in SLA has the function of a vehicle to transport learners through stations of linguistics, that skills, and activities are in ascending level of complexity, that meaning takes a priority over form and that the completion of the task is the criteria for assessment. Ellis (2003, p.16) provides the following definition of a pedagogical task in TBLT:

A task is a work plan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it

requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose a particular form.

However, some argue that form and meaning are like two wings for the language bird to fly and soar high and therefore both form and meaning should be the center of teaching and learning in SLA. In other words, they believe in the marriage between form and meaning; a marriage that is based on equal responsibilities and duties. For instance, Skehan (1998) and Long and Norris (2009) believe that task based language teaching is an attempt to harness and integrate the benefits of a focus on meaning via adoption of an analytic syllabus while simultaneously, paying attention to language learning in terms of form to deal with its unknown shortcomings, particularly rate of development and incompleteness where grammatical accuracy is concerned (Long & Norris, 2009, p. 599).

The main objective of TBLT is language acquisition, compared to sole successful communication; language learning is a priority in terms of fluency, accuracy and complexity. There are some famous models of TBLT among which is Willis's model that includes a pre-task, a task, and a post task or language focus that focuses on the language objective. The task phase is mainly engaging for the learners taking advantage from the teacher being in a passive mode. Another model is Long's model which employs real life tasks to teach the language. The models include a needs analysis to identify task targets and accordingly task types. Based on these task types, pedagogic tasks are designed and sequenced (González-Lloret, 2015). The present investigation is very close to this model as it is a task based assessment of students' speaking performance through three tasks similar to IELTS speaking test.

B. Interactional Competence

Closely relevant to the exploration intended in this study is the work of Walsh (2011), who defines interactional competence as "teachers' and learners' ability to use interaction as a tool for mediating and assisting learning" (p. 158). Walsh criticized the fact that teachers and examiners focus only on the individual performance in terms of accuracy, fluency, and lexical resource and neglect the importance of negotiating meaning and clarifying points of view; the interactive aspect of speech is not a priority.

The notion of interactional competence was first coined by Kramsch (1986) who argues that teachers' focus should be on interactional competence to enhance learners' ability of competent communication that involves mutual understanding between interlocutors.

In essence, Interactional Competence is relevant to 'Confluence', which was defined by McCarthy (2005) as the act of making spoken language fluent together with another speaker. Spoken confluence is highly relevant to the third task of this study, the mini debate which involves interaction by two interlocutors who express points of view and refute opposing ones. In accordance with this view, Markee (2008) states that interactional competence in a second language involves learners "co-construct[ing] with their interlocutors locally enacted, progressively more accurate, fluent, and complex interactional repertoires in the L2" (p. 3).

Acknowledging the role of context, Young (2008) offers the following definition of interactional competence: "a relationship between participants' employment of linguistic and interactional resources and the contexts in which they are employed" (p. 100). Due to the salient attention to interactional competence, it can be rightly predicted that this type of competence might soon mount to become a fifth skill (Walsh, 2012) added to the pool of the four known skills: speaking, reading, writing and listening.

C. Research on Task Complexity

Most of the studies conducted in this regard revealed that language learners tend to be slow when they produce language in response to tasks that are more complex (Gilbert, 2005; Rahimpour, 1997; Robinson, 1995). As far as lexical complexity is concerned, most studies show that language learners produce more advanced and varied levels of vocabulary when the language tasks are more complex. With narrative tasks, both Robinson (1995) and Gilbert (2005) exhibited an impact of increased task complexity on lexical complexity. Révész (2012) and Michel et al. (2008) confirm the same in their studies that added more elements required for reasoning like the number of features in mobile phones and the number of instructions in a task that focuses on decision making strategies.

Conversely, structural complexity almost witnessed no impact of task complexity on L2 learners' language production. Many researchers like Robinson (1995), Rahimpour (1997) and Gilbert (2005) found no significant effects for increasing task demands on structural complexity. As for Studies manipulating the number of elements in the spoken language production (Michel, Kuiken & Vedder, 2007), and in written performance (Kuiken & Vedder, 2007; Kuiken, Vedder & Mos, 2005), researchers did not find any effect of increasing cognitive complexity on structural complexity. The latter two studies manipulated the task complexity in written tasks by increasing the number of instructions given to participants. There was only one exception to these findings; Révész (2012), who found substantial impacts for increased numbers of elements and amount of reasoning on the structure of produced language.

As for accuracy, the studies found contrasting results; whereas Robinson (1995), Ishikawa (2007), Iwashita et al. (2008) and Gilbert (2005) have shown that learners tend to be more accurate, Rahimpour (1997) and Révész (2012), confirmed no effects have been detected for increasing task complexity on the accuracy of learners' production. Most of these contrasting results can be justified due to the different assessment criteria and tools in addition to the difference in setting and samplings.

Dialogical tasks, though, have received much less attention. Michel et al. (2007) investigated the increase of task complexity on learners' individual and interactive performance by increasing the number of features in a task that required a decision-making process. The findings showed that learners tended to slow down the pace of their speech and they became more accurate; their linguistic complexity was not affected. In addition, substantially higher levels of fluency and accuracy were found in the dialogical interactive tasks than the individual tasks. Robinson (2007) used a series of three narratives from the Wechsler Adult Intelligence Scale-Revised (WAIS-R) which increased in the amount of intentional reasoning. Robinson found very few effects of task complexity on production as measured by general measures, while he found a significant effect with more specific measures such as the number of psychological state terms. Generally, most of the studies related to the impact of task complexity on monologic tasks have shown a strong inclination for fluency to become less, and for lexical complexity to rise, yet no effects have been found on structural complexity. As for accuracy, the picture was not clear as mixed findings have shown. On the other hand, dialogic or interactive tasks were shown to cause fluency to increase while they caused accuracy to decrease. Both lexical and structural complexity have been influenced. The various proficiency levels of learners participating in these different studies is one main reason for the differences in findings. Another reason may be the use of one single task type with some groups while using different task types with other groups in different studies. Finally, maybe the measures used to trace the impacts of task complexity were not reliable or accurate.

III. RESEARCH METHODOLOGY

A. Approach

The researchers espoused the qualitative approach, which is very suitable for the present study as it helps to gain detailed insights based on the actual task-based speech production by L2 learners. Grounded in anthropology and sociology, the qualitative approach derives from social constructivism which contends that knowledge is constructed during interactions in the real world (Creswell 2013). Interest in this epistemological perspective led researchers to focus on the principles of qualitative methods compatible with research adopting the social constructivist philosophy. The goal of qualitative research is to explore certain human interactions in particular social settings. Marshall and Rossman (2011) recommend investigators access the real world of the participants, wherein they understand the meanings introduced by the participants to the settings during their interactions.

B. Setting and Participants

This study was conducted in an American-curriculum-based school in Dubai, United Arab Emirates. The appropriate school permission and parents' consent were obtained to conduct the study. The students were very excited to participate in this study because of the similarity between the tasks and the IELTS speaking test that they must take to graduate and enroll in universities. The participants were tested in a classroom with the presence of one of the researchers; participants were individually tested in the monolog task and in the dialog/discussion task. The 3rd task, the mini debate, was conducted with the presence of two participants once and one participant and another non participant student once.

The participants in this study were students in grade 12 from the above mentioned school; the number of students who participated in the study was 4 girls: two high achievers with an actual IELTS scores of 7.5-8.5 and two above average with scores of 6.5 each. Participants were informed that they would be asked to perform a speaking test similar to the IELTS speaking test with its three tasks: a monolog, a dialog, and a debate-like discussion. All the four participants were aware of the IELTS test procedures, and they were familiar with the test as achieving an Academic IELTS band of 5.5 was a requirement and a prerequisite to enrolling in any university. The participants were well acquainted with one of the researchers being their English teacher. This point was important as it reduced the tension or the examiner-caused stress to almost zero, hence external factors of impacting their speech were greatly eliminated and were only limited to the cognitive aspect and the linguistic complexity. For research ethics, permission was obtained from all four participants as well as their guardians. Participants showed willingness to participate and their guardians showed willingness to cooperate.

C. Instrumentation and Tools

Participants' speech and responses during the three tasks of the study were recorded and transcribed. A Sony Voice recorder (IC recorder), which had a capacity to record up to 90 minutes, was used to conduct the speaking test with its three tasks that range and ascend in complexity from a simple monolog to dialog and finally to a mini debate between two participants. The recordings were listened to several times in order to be able to evaluate each part of the speaking criteria. Special attention was given to listening to and analyzing the arguments in the 3rd task, the mini-debate in order to evaluate the reasoning competence and cognitive skills when increasing task complexity. The second data gathering instrument was semi-structured interviews. These interviews were conducted with the four participants to investigate their perceptions and attitudes towards different tasks they had to perform. Furthermore, participants' interviews after conducting the study were audio recorded, transcribed, and analyzed. Finally, the IELTS band descriptor was utilised as an assessment tool of participants' oral speech production during the three tasks of the study.

D. Research Procedure

The study was conducted in the school premises in one of the classes in the high school building. Students were examined individually for the monological tasks and the interview task. They were paired for the mini debates and they were observed, and audio recorded by one of the researchers. Later, the researcher interviewed the students and asked them about the three tasks. All responses were recorded and transcribed to be processed for analysis using the IELTS band descriptor found in the IELTS official website and handbook. Two IELTS examiners analyzed the speech responses again to insure a very objective and professional feedback and a high rate of interrater reliability.

IV. RESEARCH FINDINGS AND DISCUSSION

A. Introduction

This chapter presents all the findings and provides explanations and possible interpretations of the findings based on the participants' speech recordings and transcription throughout the three speaking tasks. The 1st task was a monolog in which participants were asked to talk about their experience in the school celebration during the national day of the UAE, December 2016. The 2nd task required participants to describe an unforgettable travel experience that they had had. Both tasks 1 and 2 are based on past experiences and require the use of past tenses. The 3rd task is a mini debate between two participants who have to argue for /against wearing school uniform; participants were given a few seconds to think of good reasons to support their arguments. The 3rd task involves more reasoning skills and it includes more cognitive complexity than tasks 1 and 2. The chapter will introduce the findings of each task following the same order they were given to students.

The monolog task is considered to be the easiest as it involves one question that requires narrating an event that was recent and participants have no pressure from an interviewer. This task is the baseline or the foundation on which the assessment of speech production will be built as the other two tasks are designed to be higher in complexity. Accordingly, this may be used as what the researcher calls a 'control task' as it involves a low level of complexity accessible to all levels of learners/participants.

In this task, as expected, the high achievers are evidently very fluent in terms of speed and the number of pauses. They are relevant and their speech production is coherent. Both high achievers spoke with a great deal of accuracy and flexibility in terms of vocab or lexical resource. Grammatical range and accuracy was another area they both scored high and their sentence structures varied to demonstrate their high proficiency level. As for mid-achievers, their level was evidently less in the monolog task as they showed less fluency by being slower and by making more pauses that are language related. Both mid-achievers used simple and average words and their structural complexity was very basic as they used mostly simple sentences and very few subordinate clauses. Also, the grammar mistakes by both participants are common in this level of proficiency unlike the high achievers who had almost no grammar mistakes. This general lower level of performance by mid-achievers naturally reflects their level and similarly it will serve as a good indicator to see how differently they perform in the other two tasks. Figure 1 shows the results of Task 1.

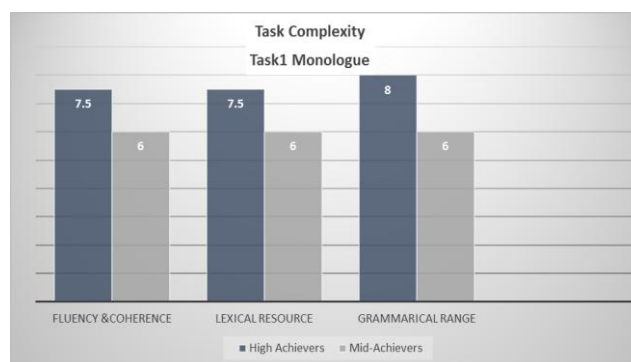


Figure 1. Results of Task 1

B. Research Question 1

The findings of the interview task show that both high achievers were still very fluent and spoke even more in terms of task duration (more than 4 minutes) compared to the monolog task (almost 3). The pausing was limited to only one occasion that is content related. Both participants used a wider range of vocabulary and more highly academic words and sets of collocations. The level of performance of the two high achievers was higher than the monolog task (their band score as per the assessment criteria rose from 7.5 to 8.5 in these two areas of speech production. Their grammatical range and accuracy was slightly affected as they produced very few errors (2 compared to 1 error in task 1) but maintaining a very high structural complexity kept their score at 8.5. These results confirm what (Gilabert, 2007; Gilabert et al., 2009) found before as high achievers naturally increased their lexical complexity level as well as their fluency; they rose to the higher complexity level. Yet, the results contradict what was suggested by (Robinson, 1995;

Rahimpour, 1997) who have confirmed that fluency of learners in spoken tasks declines if participants are asked to perform a task involving narration without looking at a visual. Finally, these findings are in line with the cognition hypothesis which states that there are likely to be positive effects on speech production when tasks are made complex.

On the other hand, both mid-achievers, were slightly affected by increasing complexity in the second task. Their fluency was a bit less as the 3rd participant's speech was slower than that of task one and she had more number of pauses. As for the 4th participant, she was not able to speak as long as the rest of the participants although she was somehow fast; their score declined from 6 to 5.5 in terms of fluency and coherence. The grammatical range and accuracy was less successful in the speech of the two mid-achievers as it was considerably affected negatively; more grammar and structure mistakes started to appear; errors were various and related to the use of countable nouns, indefinite articles, plurals, and subject-verb agreement. It seems that participants had a trade-off of skills in this area as they focused more on content and not form. The band score achieved in this area declined (from 6 in task one) to 5. The findings in this area of speech production assessed confirm Skehan's (1998) 'Limited Capacity Hypothesis' which claims that more cognitively complex tasks "consume more attentional resources...with the result that less attention is available for focus on form" (p. 97), therefore, language output like grammar rules, spelling, and punctuation can be less accurate. The results of Task 2 can be shown in Figure 2.



Figure 2. Results of Task 2

In the mini-debate task, both high achievers were evidently very fluent as their speech production was fast enough and greatly coherent; they had almost no pauses or hesitations. Both participants outperformed themselves in the previous task; they received a score of 8.5 compared to 8 in the previous tasks.

In terms of lexical resource, the impact of task complexity was similarly positive as both participants used more advanced level of academic. The score accordingly rose from 7.5 to 9 which is a remarkable progress. Both high achievers, used mostly correct grammar rules (only two errors in 9.5 minutes) and their sentences structures were varied tending to be more complex; in other words, their level of performance also rose to meet the complex demands of this task.

The fourth criteria of assessment, reasoning and argumentation added more cognitive pressure on them. Both high achievers used valid arguments. The interactive nature of dialog and the shared or divided pressure of speaking reduce the stress during language performance because speakers get turns and accordingly time to plan during their partners' turns (Tavakoli & Foster, 2008). When in a monological task, the speaker feels all the stress alone and hesitation is an element that plays a role in speech production. In dialogues, both interlocutors have higher chances of accuracy and fluency in terms of speed and repair as they both have more attentional resources available (Levelt, 1989).

On the other hand, the two mid-achievers level of performance evidently declined as their speech suffered from more unintentional pauses (around 11 times) and they barely lasted for 5 minutes compared to the high achievers who lasted for more than 9 minutes. It is also noteworthy that the 4th participant was not able to continue performing this task of debating due to mental cognitive stress as she translated that into a sentence in L1, 'my mind has been blocked'. The score fell from 6 to 5. The findings in this area of fluency and coherence agree with Skehan's (1998) 'Limited Capacity Hypothesis'. In terms of lexical resource both participants performed with the same level as the previous task; there was no clear impact on their speech production except for the scarcity of high academic vocabulary yet they still perform well as per their level. The stability of performance in this area maybe due to the nature of interactive task. Further, it can be explained in terms of what was stated by Pickering and Garrod (2004) who believe that in dialogical tasks learners borrow and reuse the structures and lexis of each other; this linguistic behavior is called alignment. As for grammatical range and accuracy, mid-achievers underperformed in this area of the task. Grammar errors were many and frequent; new errors appeared like pronoun-antecedent agreement and uneasy shifts in tenses, in addition to some wrong sentence structures that were only simple throughout the entire task. The score declined to reach 4.5, the lowest score throughout all research tasks and areas. The findings in this area contradict those of Michel et al. (2007) who investigated the increase of task complexity on learners' both individual and interactive performance and his research findings confirmed decreased fluency and increased accuracy and structural complexity.

The two mid-achievers' reasons were not enough to support and arguments were not fully refuted. The cognitive stress caused many mistakes in term of language as well as the logical sequence of the debate as the participants kept

repeating their words in some occasions during their task which lasted for only 5 minutes. In general, the evaluation of both external raters confirms the patterns found in the evaluation of the researchers. In other words, high achievers ascendingly received higher scores in terms of lexical resource across the three tasks while the mid-achievers descendingly received lower scores in terms of grammatical range and accuracy across the three tasks. The results of Task 3, the mini debate, can be shown in Table 1.

TABLE 1.
THE RESULTS OF THE MINI DEBATE

Evaluation criteria	High achievers scores	Mid-achievers scores
Fluency and coherence	8.5	5
Lexical resource	9	6.5
Grammatical range	8	4.5
Reasoning and debating skills	8	6

C. Research Question 2

Evidently, high achieving participants had a better performance in the interview and the mini debate tasks; in fact, both participants scored higher in terms of fluency, accuracy, and most importantly in lexical resource or complexity. Such a positive correlation confirms the findings (except for lexical complexity) by Michel et al. (2007) who investigated the increase of task complexity on learners' both individual and interactive performance by increasing the number of features in a task that required a decision-making process. The findings showed that learners tended to slow down the pace of their speech and they became more accurate; their linguistic complexity was not affected. In addition, substantially higher levels of fluency and accuracy were found in the dialogical interactive tasks than the individual tasks. Another researcher Gilabert (2007) found a strong correlation between participants' language proficiency and their performance in dialogical tasks. Accordingly, high achievers' performance demonstrated a noticeably positive correlation between their level and their performance in dialogical, interactional tasks.

As for the mid-achievers, the findings suggest that in dialogical tasks their performance received a negative impact and they scored less than the individual task (monolog) in terms of grammatical range and accuracy in which they scored 4.5 in the mini-debate on the IELTS band score. As for the rest of other language production areas examined, the findings were similar to (Gilabert et al., 2009) weaker correlations found between proficiency and performance in all dialogical tasks.

The positive correlation between high achieving participants' proficiency and their performance in dialogical tasks accords with the suggestions of many researchers (e.g. Schmidt, 1990; Swain, 1985), who found a positive impact of interaction in speaking tasks, dialogs. In addition to those, the Cognition Hypothesis also predicts higher accuracy measures in dialogues (Robinson, 2005, 2007). An illustration of the correlation found in this study between the participants' proficiency levels and their performance is provided in Figure 3.

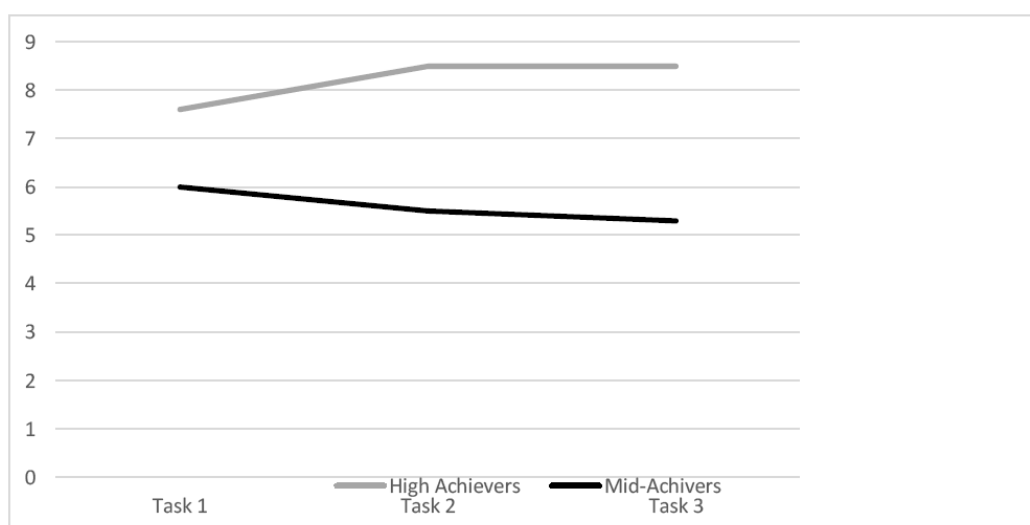


Figure 3: Proficiency Level Correlation with Performance

With regards to the external raters' evaluation, there was insignificant variation from the researchers' findings, and the overall analyses were almost the same in terms of fluency, lexical resource, and grammar. Both raters evaluated the recordings and rated them using the same evaluation criteria; The IELTS speaking band descriptor. The ratings and scores given by both examiners confirm the patterns found in the evaluation by the researcher. To elaborate, the first rater's and the second rater's evaluation of the impact of increasing complexity on lexical resource in high achievers'

production showed a steady increase in vocabulary, compared to a steady drop apropos the grammatical range and accuracy in mid-achievers' productions.

D. Research Question 3

This research question was set to find out whether or not increasing the cognitive complexity of the task would have an impact on oral language performance. The main negative impact was related to the grammatical range and accuracy as all participants' performances were affected; the influence ranged from dramatic (in the case of mid-achievers) and slight (in the case of high achievers). Nevertheless, a positive correlation was found in the lexical complexity area in which both high achievers and mid-achievers scored highest in the debate task that required more reasoning skills and underpinned more cognitive load. To some extent, the findings of this research question merit more consideration as they do not agree with the results of (Gilabert, 2007) and (Gilabert et al., 2009), which concluded that the increase of task complexity had no impact on language performance in the tasks that involved decision making or cognitive reasoning demands.

V. CONCLUSION AND RECOMMENDATIONS

A. Conclusion

This study investigated the impact of task complexity on language performance and production in ESL classrooms. Recent literature findings revealed that both negative and positive effects of increasing task complexity on language areas like fluency, lexical complexity, and grammar and accuracy. The researchers chose speaking as the skill to be put under the microscope in this study.

Links were made to the theoretical backgrounding of the present investigation. Relating the present paper to cognition hypothesis was an important one as the study examines the effects on fluency, lexical complexity, grammar and accuracy when the tasks are made more complex. Similarly, the limited capacity hypothesis was importantly relevant as it suggests that task complexity has negative effects due to the consumption of attentional resources.

The findings of the monolog task showed that the high achievers were evidently very fluent in terms of speed and the number of pauses. Their speech production was coherent. Both high achievers spoke with a great deal of accuracy and flexibility in terms of lexical resource. Grammatical range and accuracy was another area in which they scored high and their sentence structures varied to demonstrate their high proficiency level. As for the mid-achievers, their level was evidently less in the monolog task as they showed less fluency by being slower and by making more pauses. Both mid-achievers used simple and average words, and their structural complexity was very basic as they used mostly simple sentences and very few subordinate clauses. Also, the grammar mistakes by both participants were common in this level of proficiency, unlike the high achievers who had almost no grammar mistakes.

The findings of the interview tasks showed that the two high achievers still had a high level of fluency and coherence and only had a slight negative effect that was barely noticeable on the grammatical range and accuracy. The only area that witnessed a very positive impact was the lexical complexity as participants used a broader range of vocabulary and more highly academic words and sets of collocations. On the other hand, the mid-achievers demonstrated a very slight negative effect on fluency and a great negative impact on grammatical range and accuracy.

The findings of the mini-debate task, the highest in terms of complexity, showed that the high achievers were still very fluent due to their speed and owing to the low number of pauses. The positive impact on lexical resource was even higher in this task. As for the grammar, there was a slightly negative impact, but the researchers did not consider it due to the length of the debate task that lasted for 9.5 minutes. The mid-achievers demonstrated less competence in terms of fluency as they spoke less time full of pauses and could not finish the debate. Evidently, the area that witnessed a major negative impact was the grammatical range and accuracy.

B. Recommendations

Based on the findings of the present study, a host of recommendations can be made. Increasing the cognitive demands of tasks merits consideration as part of SLA classroom practices that involve focus on form as increasing the cognitive load demonstrated the direction of learners' attention to form, either to the vocabulary they need to use or to the structures they need to consider. In addition, data about how increasing cognitive complexity impacted performance should inform decisions on how to sequence tasks in a syllabus. Sequencing tasks in an ascending order could be handy as the performance of learners rises gradually. Another important point is that interactive tasks like interviews and debates push L2 learners to greater accuracy, lexical complexity, and fluency. Speaking activities or tasks therefore should incorporate pair work or group work to trigger more fluency and lexical complexity. Language teachers and testers however should be aware that L2 learners are often evaluated in a monologic setting, especially in a testing environment. This is likely to mislead/distort the evaluation of learners as they can perform better in a dialogical testing setting. Finally, teachers, and syllabus designers can make use of task manipulation to adjust the level of tasks and instructions to suit the needs and levels of learners.

REFERENCES

- [1] Baralt, M. & J. M. Gomez. (2017). Task-based language teaching online: A guide for teachers. *Language Learning & Technology*, 21.3, 28-43.
- [2] Batstone, R. (2016). Rethinking TBLT in context: The negotiation of form. *The Language Learning Journal*, 44.4, 506-518. "doi:10.1080/09571736.2016.1227221".
- [3] Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research* 3.3, 185–214. "doi:10.1177/136216889900300302".
- [4] Cadierno, T. & P. Robinson. (2009). Language typology, task complexity and the acquisition of L2 lexicalization patterns for reference to motion. *Annual Review of Cognitive Linguistics* 7.1, 245–276. "doi:10.1075/arcl.7.10cad".
- [5] Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *The Modern Language Journal*, 102.1, 162-180. "doi:10.1111/modl.12460".
- [6] Costa, A., M. J. Pickering & A. Sorace. (2008). Alignment in second language dialogue. *Language and Cognitive Processes* 23.4, 528–556. "doi:10.1080/01690960801920545".
- [7] Cresswell, J. W. (2013). *Research design: Qualitative, quantitative and mixed methods approaches* (4th edn.). Thousand Oaks: Sage.
- [8] Gagne, R. M. (1962). Military training and the principles of learning. *American Psychologist* 17.2, 83–91. "doi:10.1037/h0048613".
- [9] Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching*, 45.3, 215-240. "doi:10.1515/IRAL.2007.010".
- [10] Gilabert, R. (2005). Task complexity and L2 narrative oral production. Ph.D dissertation, University of Barcelona.
- [11] Gilabert, R. (2007). The simultaneous manipulation along the planning time and \pm Here-and-Now dimensions: Effects on oral L2 production. In M. D. P. G. Mayo (ed.), *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 44-68.
- [12] Gilabert, R., J. Barón & A. Llanes. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics*, 47.2-3, 367–395. "doi:10.1515/iral.2009.016".
- [13] González-Lloret, M. (2015). *A practical guide to integrating technology into task-based language teaching*. Washington, D. C.: Georgetown University Press.
- [14] Hyun, J. & J. Lee. (2018). The effects of task complexity and working memory on Korean adult learners' English speaking performance. *English Teaching*, 73.1, 115-134. "doi:10.15858/engtea.73.1.201803.115".
- [15] Ishikawa, T. (2007). The effect of increasing task complexity along the [\pm Here-and-Now] dimension on L2 written narrative discourse. In M. D. P. G. Mayo (ed.), *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 136156.
- [16] Iwashita, N., T. McNamara & C. Elder. (2008). Can we predict task difficulty in an oral proficiency test?: Exploring the potential of an information processing approach to task design. *Language Learning*, 51.3, 401–436. "doi:10.1111/0023-8333.00160".
- [17] Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [18] Kramsch, C. (1986). From Language Proficiency to Interactional Competence. *The Modern Language Journal*, 70.4, 366-372. "doi:10.1111/j.1540-4781.1986.tb05291.x".
- [19] Kuiken, F. & I. Vedder. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. D. P. G. Mayo (ed.), *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 117-135.
- [20] Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- [21] Long, M. H., & Norris, J. M. (2009). Task-based language teaching and assessment. In K. V. D. Branden, M. Bygate & J. M. Norris (eds.), *Task-based language teaching: A reader*. Amsterdam: John Benjamins, 135-142. "doi:10.1075/tblt.1.09tas"
- [22] Markee, N. (2008). Toward a learning behavior tracking methodology for CA-for-SLA. *Applied Linguistics*, 29.3, 404-427. "doi:10.1093/applin/amm052".
- [23] Marshall, C. & G. B. Rossman. (2011). *Designing qualitative research* (5th edn.). Thousand Oaks, CA: Sage.
- [24] McCarthy, M. J. (2005). Fluency and confluence: What fluent speakers do. *The Language Teacher*, 29.6, 26-28.
- [25] Michel, M., F. Kuiken & I. Vedder. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics*, 45.3, 241–259. "doi:10.1515/iral.2007.011".
- [26] Rahimpour, M. (1997). Task condition, task complexity and variation in L2 discourse. Ph.D. dissertation, University of Queensland.
- [27] Reigeluth, C. (1999). The elaboration theory: Guidance for scope and sequence. In C. Reigeluth (ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates, 425-452.
- [28] R é v é s z, A. (2012). Task complexity, focus on form-meaning connections, and individual differences: A classroom-based study. *Modern Language Journal*, 95.s1, 162-181. "doi:10.1111/j.1540-4871.2011.01241.x".
- [29] Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43.1, 1–32. "doi:10.1515/iral.2005.43.1.1".
- [30] Robinson, P. (2011). *Second language task complexity*. Amsterdam: John Benjamins Publishing.
- [31] Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45.1, 99–145. "doi:10.1111/j.1467.1770.1995.tb00964.x".
- [32] Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 43.3, 1–31. "doi:10.1515/iral.2007.009".
- [33] Seedhouse, P. (2005). Conversation Analysis as Research Methodology. In K. Richards & P. Seedhouse (eds), *Applying conversation analysis*. Basingstoke: Palgrave Macmillan, 251-266.
- [34] Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- [35] Skehan, P. (2002). A non-marginal role for tasks. *ELT Journal*, 56.3, 289–295.
- [36] Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36.1, 1–14. "doi:10.1017/S026144480200188X".

- [37] Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (eds.), *Input and second language acquisition*. Rowley, MA: Newbury House, 64-81.
- [38] Svennevig, J. (2003). Echo answers in native/non-native interaction. *Pragmatics*, 13.2, 285–310. “doi:10.1075/prag.13.2.04sve”.
- [39] Tavakoli, P. & P. Foster. (2008). Task design and L2 performance. *Language Learning*, 58.2, 429–473.
- [40] Van den Branden, K. (2006). *Task-based language education: From theory to practice*. Cambridge: Cambridge University Press.
- [41] Van den Branden, K., M. Bygate & J. Norris. (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamins.
- [42] Vasylets, O., R. Gilabert & R. M. Manchon. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67.2, 394-430. “doi:10.1111/lang.12228”.
- [43] Young, R. (2008). *Language and Interaction: An Advanced Resource Book*. London: Routledge.



Amir A. Hassanein was born in Al Ta'ef, Saudi Arabia in 1980. Hassanien achieved his M.A. degree from the British University in Dubai in Education/TESOL in 2017, and his B.A. degree in English literature and translation from Faculty of Al -Alsun (Languages), Ain Shams university in Cairo, Egypt in 2001.

He currently holds the position of research assistant in the faculty of education in the British University in Dubai. He also holds the position of an English lecturer in the University of Modern Sciences, Dubai and as a grade 12 senior English teacher in Dubai International school, United Arab Emirates. He also worked as an English instructor in Al Dar University College and a translator for different entities. His research interests include teaching methods, sociolinguistics, language education, corpus linguistics and literary analysis

Mr. Hassanein was an organizing committee member of TESOL Arabia, Dubai chapter. He was also a member in the accreditation committee of AdvancED in Dubai International School.



Emad A. S. Abu-Ayyash was born in Irbid, Jordan in 1974. Abu-Ayyash achieved his Ph.D degree in Education/TESOL from the British University in Dubai, United Arab Emirates in 2016, his M.A. degree in translation from Yarmouk University in Irbid, Jordan in 2007 and his B.A. degree in English literature and linguistics from Yarmouk University in Irbid, Jordan in 1996.

He currently holds the position of assistant professor in the faculty of education at the British University in Dubai. He also worked as an Instructional Leadership Coordinator, English lecturer, English Team Leader, English Teacher and Translator. His research interests include discourse analysis, teaching and learning, translation, assessment, and TESOL.

Dr. Abu-Ayyash is an associate member of GEN TEFL International Conference, scientific committee member of Elsevier's ACLing Conference on Arabic Computational Linguistics and an editorial board member of Journal for Researching Education Practice and Theory.