

The Impact of Consciousness Raising and Communicative Tasks on Fluency and Accuracy during Interaction within Saudi Secondary School Students

Adnan Mukhrib

English Language Centre, Umm Al Qur'a University, Makkah, Saudi Arabia

Abstract—This study investigated the effect of collaborative tasks on language performance, with a focus on fluency and accuracy. Three groups of Saudi learners of English in a secondary school were given either awareness raising tasks, meaning-based activities or a combination. Quantitative data collected from 72 Saudi learners of English and then analysed to answer the research questions. The test results were analysed for speech fluency, written fluency, accuracy and lexical richness to provide quantitative measures of any improvement over the three test periods. Consideration was also given to language related episodes (LREs), i.e. pauses, repetitions and self-correction to identify changes during the interactions. The results indicated that there was an improvement in both fluency and accuracy. In addition, there is a likely significant benefit of encouraging interaction amongst peers during a task-based learning approach, particularly when there is a combination of CT and CR tasks, when compared to one task type. The findings underline the relationship between classroom interaction and practice and improvements in fluency and accuracy. This is important due to the increasing view that in the teaching of EFL, fluency is being neglected despite its importance in achieving communicative competence as Tavakoli and Hunter (2018) noted. As a result of the findings we argue that Saudi secondary school EFL classes should place a greater emphasis on TBL and interaction, combining both consciousness raising and communicative tasks to maximise the improved benefits seen in this study.

Index Terms—consciousness-raising, communicative tasks, interaction, fluency, accuracy

I. INTRODUCTION

The benefits of collaborative learning, i.e. interaction, are considered to be a fostering of social connections and encouragement of student engagement in class activities (Han, 2014; Shehadeh, 2011). These identified benefits have contributed to the rise in adoption of a task-based learning (TBL) approach although there is a need to determine whether consciousness raising (CRT) or communication tasks (CT) deliver the most promising outcomes for students. Certainly, problem solving tasks are considered to be a highly effective way to introduce and encourage collaboration and interaction leading to identification and correction of errors that might not occur in a more passive teaching environment. We therefore believe that introducing the idea and practice of collaborative learning into the Saudi school environment, will demonstrate the benefits to learners of both using CT and CRT tasks.

This study examined learner outcomes using these CRT and CT tasks with Saudi learners of English. We first provide a theoretical and methodological overview of previous research into task-based learning, interaction, fluency and language related episodes. This is followed by a description of our research procedures and a presentation of the findings. Next, we discuss the results of our research with reference to psycholinguistic processes of L2 speech production and perception. We conclude our paper by highlighting the benefits of the task types and the interaction and outlining future directions for research.

II. LITERATURE REVIEW

A. *The Nature of Collaborative Interactions*

Research into task-based learning suggests that when students work collaboratively, each individual learner brings their own knowledge and perceptions and speaking style into the group meaning there may be divergent views and various levels of knowledge and interaction styles (de Jong et al., 2015). The interaction is a convergence of these factors which enable sharing of knowledge and collaboration for task resolution. Again, this reflects on the process of learning as a social practice, and may involve a number of strategies related to students' own personalities and speaking styles. These include assertiveness, or domination of others, acquiescence from silent students or more passive members of the group. Puntambekar (2006) suggests an ideal collaborative group make-up is one where there is reciprocal sense making of the task and its required outcomes and components. This entails moving from disagreement to assimilation

and finally to the construction of new understanding through discussion and negotiation about the task and its requirements. Longer term, the aim is that the shared knowledge creation increases with familiarity amongst group members and the generation of new understanding of the requirements of the L1 (Hmelo-Silver, Chernobilsky & Jordan, 2008).

This development of shared understanding and increased knowledge is reflected in the language produced during and after the collaborative process. Of particular interest is a potential conflict between the cognitive approach and the knowledge the individual students have in their heads, and the group sharing and interaction (Elliott & Zhang, 2019). However, when there is a need to consider developing both implicit and explicit knowledge of how language works and thus the ability to use it for a problem solving task; this conflict is not necessarily a major issue. The process of top down and bottom up learning means that whichever route is taken there is a need to match the linguistic knowledge and speaking style of the individual to contextual and situational knowledge and real world factual information regarding how language is used and thus applied to task resolution. When the cognitive and interactionist approaches are used together, a two-way link is created that enhances, through collaboration, the overall learning process and language production (Hinkel, 2006).

B. Collaborative Language Production

In terms of the language produced through the interaction, the intent is to encourage students to focus on the form of what is produced as a group. The collaborative nature of the task encourages group and individual awareness of knowledge gaps, that may be filled by shared information via discussion and exploration of forms (Nassaji & Tian, 2014). Furthermore, the discussions, collaborations and knowledge sharing increase awareness of the relationships that exist between function, form and word meaning during interaction in the TL, enhancing students metacognitive understanding of the nature of the L2 (Jong, 2009; Leiser, 2004). Moreover, the peer feedback during collaboration creates a cyclical approach to examining language knowledge and production, this in turn promotes the deeper understanding necessary for language proficiency. The view is that producing language collaboratively, rather than individually supports understanding of how language learning works, in a natural interactive way that is more reflective of real world situations (Kowal & Swain, 1994). This however is dependent on the effectiveness of the interactions, group dynamics and whether the speaking styles are complementary according to Trofimovich et al. (2019).

C. Group Interaction and Dynamics

When there is a positive group dynamic, there is concurrent enhanced motivation, moral and ultimately learning outcomes (Gorse & Sanderson, 2007), as the group shares goals and responsibility. Achieving this optimum team can however be a challenge in an EFL classroom where there may be diverse personalities, knowledge and even proficiency levels (Zastrow, 2009). There is a need to create a sense of belonging which contributes to a rewarding learning experience and encourages students to work together (Gorse & Sanderson, 2007). Following the work of Belbin (2010) who identified specific roles within a group, in this study it was recognised that good practice would require initial team building, despite a level of existing familiarity between the students. It was evident all the different role styles were present (planters for ideas generation, resource investigators for solution location and conflict resolution), all of whom contributed to effective negotiation of meaning (Gass et al., 2011; Long, 1996; McDonough, 2004).

Alongside identification and understanding of group dynamics, this study was informed by information processing theories and the role of interaction in L2 acquisition. In essence, whether CT or CR tasks had the most significant impact on fluency and interaction, along with the influence of language-related episodes (Block, 2003; De la Fuente, 2002; Ellis & He, 1999; Swain, 2013).

D. Communicative and Consciousness-raising Tasks

Mohamed (2004) noted that there are comparable results for both CT and CRT tasks on different dimensions of learning and development. It appears that CR tasks facilitate the teaching of linguistic features, their long-term retention, and that awareness is vital for this to be effective (Van Lier, 2001). The reflection on language that results from greater awareness draws attention to decision making in L2 production and how to use this knowledge for task completion (Fotos, 1994; Izumi, 2002; Nassaji & Tian, 2010). The challenge with taking this view is that little consideration is given to the automatization of processes, deployed when learning chunks of language by rote (Wray, 2002), which is common in the completion of communicative task. Despite this, as Jeon and Hahn (2006) point out, CR task appears to be supported by collaborative learning, and thus incorporating this task type into the study has merit and value. A wider discussion of the two task types is beyond the scope of this work but further detailed information can be found in studies by Ellis (2002) and Nunan (1991).

E. Research Question

The main focus for the current study was “What is the effect of interaction on fluency and accuracy when used as part of a TBL approach? From this aim, three sub-questions emerged:

RQ 1a) How does interaction during communicative tasks influence fluency and accuracy?

RQ 1b) How does interaction during consciousness-raising tasks influence fluency and accuracy?

RQ 1c) How does interaction during combination of tasks influence fluency and accuracy?

We hypothesised that interaction during both CT and CRT tasks improved fluency and accuracy due to the shared negotiation of meaning that occurs during collaborative tasks of either nature. At the same time, we further hypothesised that based on the view that learning was a personal experience that there could be an influence from social and group dynamics on the achievement and retention of any improvement in fluency and accuracy.

III. RESEARCH METHODOLOGY

A. Setting and Participants

This research was carried out with the use of the four EFL intact classes at a secondary school in Makkah, Saudi Arabia. There were a total of 116 adult learners of English recruited for the study, aged between 17 and 18, but only 72 learners who completed all three tests are included in the results. The classes were divided into the following groups: a CT group ($n=16$), CRT group ($n=20$), combined group ($n=16$), and control group ($n=20$). They were recruited as volunteers, and all studied English at grade three as part of their overall curriculum. Arabic is their mother tongue and they had studied English language for five years and were currently in mandatory English classes at their school.

B. Approach and Procedure

The approach adopted involved an intervention in a secondary school, and the use of both CT and CRT tasks as a new means of instruction. Outcomes were assessed with tests of a quantitative nature. A total of four groups were created: Group A undertook consciousness raising (CRT) tasks, Group B performed communicative tasks (CTs), Group C undertook a combination of CT and CRT tasks and a control group just completed the tests. All the participants undertook an English Placement Test (EPT) to identify their proficiency levels and to ensure mixed abilities in the group so that interactions could be observed in terms of negotiation, support and speaking styles, and the influence of more proficient students on those less knowledgeable during the task process.

The English participial adjective was selected as the target structure because it had not yet been taught at the commencement of this study. As such it was unfamiliar to all the groups and meant that it was possible to investigate whether the inclusion of the different task types affected the use of form and ability to learn the structure. In the CRT group, the students were asked to work together to identify differences in meaning and structure for ‘-ed’ and ‘-ing’ adjectives, before formulating a rule for the use of participles, followed by making sentences to focus on the accuracy (see Appendix A). For the CT group the task was to discuss pictures and then tell stories using a number of supplied adjectives (see Appendix B for this task). The aim was to focus conveying meaning, but without full explanation of information that would ensure target-like proficiency in order for them to identify gaps in their own knowledge. At the same time, in line with findings from Byrd (2005), it was anticipated that the learners would be fluent but at the expense of accuracy. However, in the combined condition, learners first completed CRTs to raise their consciousness of the same structure. Subsequent tasks were meaning-based to assist the students to integrate the form in order to develop their interlanguage system. The control group process continued with normal lessons, and taking of the tests using the same intervals for pre-post and delayed tests as the experimental groups.

The three experimental groups attended five 45-minute treatment sessions where they performed their allotted tasks involving the use of participial adjectives. Before engaging with the task completion, each group was given five minutes planning time at all the test stages. The most straightforward approach therefore was to teach according to the regular course book in current use, it was only the mode of teaching that was adjusted. Informed consent was obtained from all the cohort, and ethical issues adhered to, as well as conforming to requirements for anonymity and confidentiality.

C. Research Instruments

The main instruments were a grammaticality judgement test (GJT), a jumbled sentence test (JST), oral production test (picture description) and written production test (for/against composition). The data gathered and the instruments used are summarised in Table 1.

TABLE 1.
TASK CONDITIONS

Groups	N	Operationalisation
CT	16	English Placement Test GJT and JST Tests Oral Production Written Production
CRT	20	English Placement Test GJT and JST Tests Oral Production Written Production
CT & CRT	16	English Placement Test GJT and JST Tests Oral Production Written Production
Control	20	English Placement Test GJT and JST Oral Production Written Production

The GJT had 20 target items and 10 distractors i.e. grammatical and ungrammatical sentences. During the test the learners were asked to identify whether the sentences were correct or incorrect based on assessment of the target structure of participial adjectives, which was unfamiliar for the students. There was no time limit set on the GJT. The JST also consisted of 20 items that the students were instructed to place the presented sentences in the correct order. The time limit was given for this task with sentence length ranging from 6 to 10 words.

In addition, the GJT was selected to provide a baseline of explicit knowledge, despite concerns raised about the test in terms of judgement making and reliability (Ellis, 1991; Han and Ellis, 1998). To manage these concerns, the JST, which is widely used for examining word order knowledge, was also applied requiring students to put 20 sentences containing jumbled words into the correct order.

D. Analysis of the Instruments Following Usage

Specific instruments used for analysis of data were a grammatical judgment test (GJT) and jumbled sentence test (JST). Although other tests (oral and written production) were utilised, these were not from specific instruments but developed from available curriculum materials and analysed using the approaches shown in Table 2 below.

TABLE 2.
ANALYTICAL APPROACHES

Item for Analysis	Analytical approaches
Speech fluency	- Number and length of pauses - Number of times a word, phrase is repeated or self-corrected are counted - Number of syllables is counted and divided by speaking time after the L1 and dysfluency have been excluded.
Written fluency	- Number of words in a text - Number of clauses - Words to total number of clauses
Accuracy	- The changes refer to syntax, word order, morphology or lexical choice. - Number of error-free clauses
Lexical richness	- Number of different words - Total number of content words

An oral production test requiring the telling of a story based on provided words and pictures, both individually and as a collaborative group provided a further level of assessment. The students were audio-recorded, with consent, to provide data for analysis in terms of fluency and accuracy, with the aim being to identify the use of effective syntax, expressiveness, coherence and fluency and whether these abilities improved for each of the groups at the three testing phases. These approaches follow the work of by Lu (2012) and Saito et al., (2016) and their work on tasks, fluency and complexity of utterances.

A final test was written production, to identify potential transfer of speech fluency and accuracy to the written word. The groups wrote a short text, collaboratively produced following evaluation of “for” and “against” information that was provided about the topic. The scale for measurement of the oral and written output was as follows:

- 1) Outputs (oral or written) were word by word. Infrequent two or three-word phrases but without contributing to meaningful syntax
- 2) Outputs (oral or written) were predominantly two-word phrase structures, awkward word groupings and unrelated to overall passage
- 3) Outputs (oral or written) were predominantly four-word phrasing, mostly appropriate word groupings and preservation of syntax is seen. However, there is limited expressive interpretation or addition
- 4) Outputs are in large meaningful phrase groups, consistent syntactic use, and a good level of expressive interpretation and additions.

This scale was adopted to provide a level of consistency in the analysis and demonstrated the skill level of the students during each testing phase. Lexical richness/density was assessed using a type/token ratio approach whilst coherence was based on logical sequencing of ideas based on the following:

- 1) No logical sequencing and awkward syntax
- 2) Minimal sequencing of a logical nature
- 3) Some logical sequencing of ideas and use of connectives
- 4) Effective, accurate use of sequencing and use of connective to provide coherence.

Taking this approach meant that for both fluency and coherence there was a potential total score of 4, and a separate lexical richness score for each test phase. Although the work was focused on the impact of interaction on fluency and accuracy, it was considered important to also examine L2 production which was assessed by examination of the grammar and jumbled sentence scores, along with speech fluency measures. Speech fluency was assessed through number and length of pauses, repetitions and self-correction as well as the speed of speech through syllables per minutes. Written fluency was identified by number of total words and clauses. Accuracy was assessed through a count of reformulations, error identification and the number of error free AS-units. From these analysis approaches, it was

anticipated that a clear indication of how the different task types influenced fluency and accuracy and what role the interaction played in achieving improvements in lexical content.

IV. RESEARCH FINDINGS

The focus of this work was on determining the effect of interaction on fluency and interaction during different types of task-based learning interventions. To identify the significant levels, ANOVA was used to analyse the gain scores in the pre-, post- and delayed post-tests obtained by the three groups. In order to answer the core research question which was “*what is the effect of interaction on fluency and accuracy when used as part of a TBL approach*”, we examined the relationship between test outcomes and interaction processes. This was combined with assessment of the level of improvement. It should be noted that only those students who completed all three tests are included in these results.

A. Grammaticality Judgment Test (GJT)

Assessment of the GJT scores by group using ANOVA analysis indicated improvement in all groups at the post-test stage. However, by delayed test stage, this improvement had reduced, although had not returned to pre-test levels. However, the CT and CR groups showed an improvement at the delayed test when compared to the pre-test.

Mean scores by group, in Table 3, indicated the greatest pre-test/post-test improvement in the control group, but again this was not retained at delayed stage, suggesting a temporary improvement. The CRT group showed the largest long term gain, indicating an increase in accuracy, which could suggest that the interactions, which were focused on raising awareness may be a contributory factor to the improvement.

TABLE 3.
DESCRIPTIVE STATISTICS BY GROUP

		Descriptive							
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
GJT Pre-test	CT & CRT	16	11.6250	4.09675	1.02419	9.4420	13.8080	4.00	19.00
	CT	16	15.5333	3.52272	.90956	13.5825	17.4841	12.00	23.00
	CRT	20	14.4500	4.98920	1.11562	12.1150	16.7850	8.00	25.00
	Control group	20	14.3500	4.68227	1.04699	12.1586	16.5414	9.00	25.00
	Total	72	14.0141	4.54656	.53958	12.9379	15.0902	4.00	25.00
GJT post-test	CT & CRT	16	17.7500	3.75056	.93764	15.7515	19.7485	10.00	25.00
	CT	16	14.1333	4.80872	1.24161	11.4704	16.7963	8.00	25.00
	CRT	20	18.1000	5.20020	1.16280	15.6662	20.5338	10.00	30.00
	Control group	20	18.3500	4.96594	1.11042	16.0259	20.6741	10.00	30.00
	Total	72	17.2535	4.93593	.58579	16.0852	18.4218	8.00	30.00
GJT delayed	CT & CRT	16	12.7500	4.59710	1.14928	10.3004	15.1996	6.00	21.00
	CT	16	15.0667	5.22995	1.35037	12.1704	17.9629	9.00	27.00
	CRT	20	16.5000	6.21120	1.38887	13.5931	19.4069	4.00	27.00
	Control group	20	12.0500	5.27631	1.17982	9.5806	14.5194	2.00	25.00
	Total	72	14.0986	5.61161	.66598	12.7703	15.4268	2.00	27.00

As the table indicates all groups showed improvement at the post-test stage except the CT group, although the combined group and control group did not maintain their improvements to the delayed stage. Conversely, the CT and CR groups both showed an improvement at the delayed test when compared to the pre-test (see Figure 1).

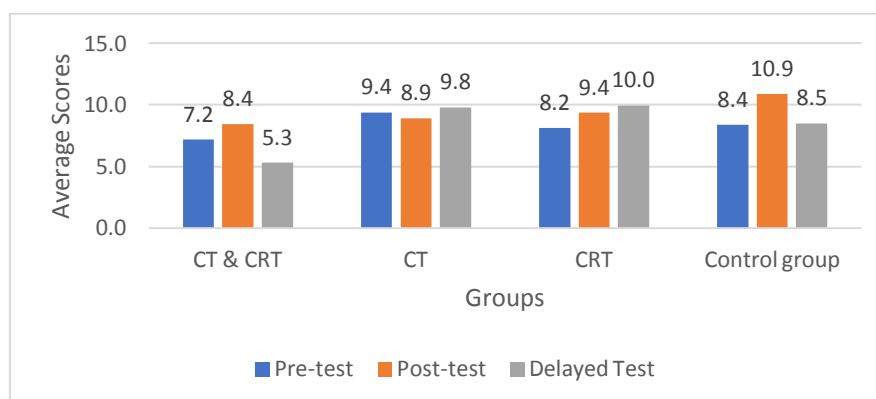


Figure 1. GJT Average Scores

To identify a correlation between intervention type and overall scores, regression analysis was undertaken which resulted in a significant relationship (.174) between the intervention type (dependent variable) and improvement on GJT scores. However, from the model summary, it was determined that only 26.6% of the improvement was due to the

intervention (CT/CRT or combined) type. In other words, the intervention only did not lead to the improvement, which again indicates a potential role for the interaction that was required to complete the task.

B. Jumbled Sentence Test (JST)

For the JST the same analysis approach was adopted, identifying initial gains at post-test but not maintained at the delayed stage. There was however variation in the overall level of significance of the improved scores for each test condition. The gradual improvement seen in all groups, suggests that the interventions are effective, but that there are other factors which may also be influencing the outcomes, for example the level of interaction in the group. Reviewing the average scores for the JST, as shown in Table 4, the CT group showed the greatest improvement, confirmed by regression and ANOVA.

TABLE 4.
DESCRIPTIVE STATISTICS BY GROUP

		Descriptive							
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
JST pre-test	CT & CRT	16	7.1875	3.25000	.81250	5.4557	8.9193	4.00	14.00
	CT	16	9.4667	5.76773	1.48922	6.2726	12.6607	1.00	19.00
	CRT	20	8.1500	5.15318	1.15229	5.7382	10.5618	1.00	18.00
	Control group	20	8.4000	4.67243	1.04479	6.2132	10.5868	1.00	18.00
	Total	72	8.2817	4.75750	.56461	7.1556	9.4078	1.00	19.00
JST post-test	CT & CRT	16	8.4375	3.98278	.99569	6.3152	10.5598	2.00	14.00
	CT	16	8.9333	6.46382	1.66895	5.3538	12.5129	1.00	17.00
	CRT	20	9.3500	5.86044	1.31043	6.6072	12.0928	2.00	20.00
	Control group	20	10.9000	4.84388	1.08312	8.6330	13.1670	6.00	20.00
	Total	72	9.4930	5.32346	.63178	8.2329	10.7530	1.00	20.00
JST delayed test	CT & CRT	16	5.3750	3.28380	.82095	3.6252	7.1248	1.00	12.00
	CT	16	9.8667	6.03403	1.55798	6.5251	13.2082	2.00	19.00
	CRT	20	9.9500	5.53910	1.23858	7.3576	12.5424	2.00	20.00
	Control group	20	8.5000	3.41051	.76261	6.9038	10.0962	4.00	15.00
	Total	72	8.4930	4.93058	.58515	7.3259	9.6600	1.00	20.00

It appears that the combined group showed the least improvement and there was in fact a loss of ability in terms of JST scores at the delayed stage. The CRT group showed the greatest improvement (1.53 at the post-stage) in the experimental groups, although surprisingly the control group showed an improvement of 2.4 average points (see Figure 2). At the delayed test stage the CT condition had the largest improvement from pre-test outcomes. Whilst these are encouraging results, it should be noted that the delayed test scores for all conditions may have been influenced by other factors during the period between the intervention and delayed testing.

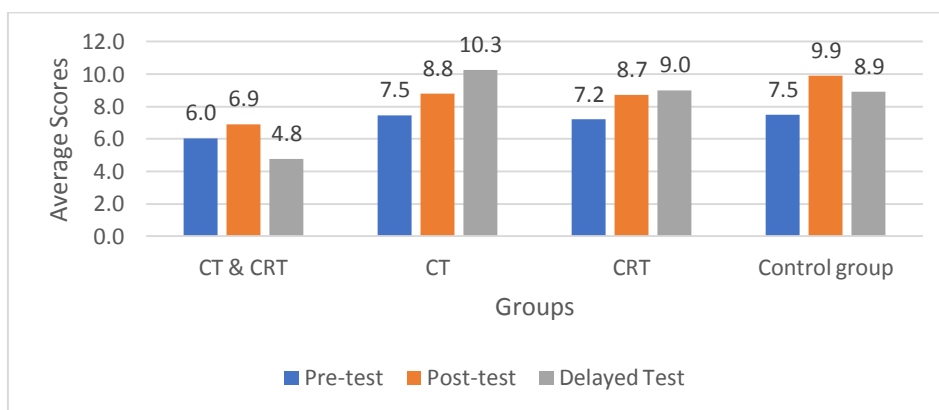


Figure 2. JST Average Scores

The suggestion overall therefore is that CT tasks provide a greater improvement on JST tests, with a 2.8% average score increase by the delayed stage from pre-test levels, potentially due to the cognitive processes involved. To confirm this finding ANOVA and regression were undertaken. However, the impact does not appear to be as significant (.124) as that seen for grammatical judgement.

C. Oral Production Task

The analysis of the oral production task was based against the scales indicated in the methodology and evaluated as percentages to facilitate between group measures. The results shown are from group, rather than individual scores, as this was felt to be a better way to measure the impact of interactions between the group members in terms of improving overall oral production. The outputs for assessment of changes in speech fluency during the oral production task were

taken from the results of each group as a whole (i.e. CT, CRT or combined and control) rather than individual student outcomes. The measures came from evaluating fluency through three sub-dimensions of breakdown, repairs and speed. Accuracy data come from application of the Foster and Wigglesworth (2016) ratios indicated in the methods section. Lexical richness data are related to the number and complexity of content words identified, through a manual count in the output. The total scores were then aggregated into the percentages shown in Table 5. The percentages refer to the level of fluency calculated based on breakdown, repairs and speed with the type-token ratio of the overall production from the group in their oral presentation. In other words, returning to the research questions, in respect of the sub-research questions, it can be seen from the results that in all cases there was improvement whilst interaction was taking place but that this was not maintained to the post test stage. In effect, the interaction appeared to improve fluency, lexical richness and accuracy for the CRT group at post-test stage (RQ1b), but only fluency and lexical richness for the CT group (RQ1a), and fluency, lexical richness and accuracy in the combined group. It is also notable that the control group who showed similar scores at the pre-test stage, and some minor improvement at post-test, did not have the same level of retention of knowledge at the delayed test stage. This suggests that even without the intervention there was some natural learning and improvement. However, the sustained improvement in the experimental groups does demonstrate the benefits of the task-based interaction and collaboration.

TABLE 5.
ORAL OUTPUTS BY GROUP

Group		Pre-test	Post-test	Delayed test
CRT	Fluency	55%	56%	54%
	Lexical richness	70%	73%	72%
	Accuracy	53%	55%	53.5%
CT	Fluency	47.5%	49%	50%
	Lexical richness	27%	54%	42%
	Accuracy	48%	47%	46%
CR & CT	Fluency	56%	57%	59%
	Lexical richness	68%	69%	67%
	Accuracy	50.5%	51%	50.2%
Control	Fluency	55%	56%	53%
	Lexical richness	53%	63%	60%
	Accuracy	51%	51%	49.3%

In evaluating core measures of fluency and accuracy, the make-up of the groups in terms of participants was separated into levels of proficiency as shown in Table 6 below (core measures of fluency and accuracy). The figures relate to the total number of pauses amongst the students in each group.

TABLE 6.
CORE MEASURES OF FLUENCY AND ACCURACY

Group	Proficiency	Participants	No. of silent pauses	filled pauses	Length of pauses	No. of repetition	No. of correction	reformulations
CRT	1	5	17	5	4	6	4	5
	2	5	12	4	4	5	1	3
	3	5	6	1	3	4	5	9
	4	5	4	0	2	0	0	2
CT	1	4	6	0	2	7	4	10
	2	4	12	5	2	5	2	13
	3	4	2	4	3	2	0	5
	4	4	9	3	3	4	0	5
CRT & CT	1	4	16	6	1	9	1	11
	2	4	13	4	1	7	4	8
	3	4	7	4	0	1	0	11
	4	4	3	1	1	2	1	3
Control	1	5	13	4	2	5	3	3
	2	5	12	4	2	6	1	3
	3	5	4	3	2	3	1	4
	4	5	5	2	2	2	1	3

Note: 1 = higher proficiency student, 2 & 3 = intermediate student and 4 = low-level student

TABLE 7.
RESPONSE TO SUB-QUESTIONS

	No. of silent pauses	filled pauses	Length of pauses	No. of repetition	No. of correction	Reformulations
CRT	39	10	3.24	15	10	19
CT	29	12	2.50	18	6	33
CRT & CT	39	15	0.57	19	6	33
Control	34	13	0.2	16	6	13

During interaction it appears that there is a short-term improvement in fluency and accuracy irrespective of task type, but that if the interaction does not continue, the improvement is lost. Of further note, and in line with work by Abu-Ayyash (2018) and Hassanein, the higher proficiency learners in all the groups had a higher percentage of reformulations, suggesting that the complexity of the task was influenced by proficiency. A similar indication can be seen from examination of the number of corrections in the different proficiency groups. Comparing the test outcomes to the control group, there was a much lower level of reformulations and corrections, suggesting again that although there was some natural learning, this was not sustained with traditional teaching. This suggests that without the interaction, the control group were unable to notice gaps and thus reformulate and correct their outputs.

The data for the scores was achieved through analysis of the discussion within each group as they performed the tasks. From this analysis, there is an indication that one potential reason for the lack of retention may be due to the make-up of the groups, and as with the grammatical tests, the discontinuation of regular interaction interrupted the retention of accuracy. This is in line with work from Spada and Lightbown (1993) who identified that when there is regular input and interaction, fluency and accuracy improvements noted during the interaction, continue to be maintained and reinforces the potential role of interaction in improving overall knowledge, fluency and accuracy during TBL tasks of both a CT and CRT nature.

D. Written Production Task

As with the oral production, the results presented are those from the group, rather than individual outputs, and was based on a 4-tier grading, and the evaluation of accuracy cohesion and lexical richness, represented by an overall percentage to facilitate comparison. As with the oral output the measures and outcomes of the group output are based on accuracy measures using the Foster and Wigglesworth (2016) ratios shown in Table 8 (Group written outputs), and lexical data were counted based on number and complexity of content words. Fluency is based on the cohesion and sentence length of the written output. Thus returning again to the sub-questions it can be seen that in relation to RQ1a written fluency, lexical richness and accuracy increased by more than 5% after the interaction for the CT group. This highlights the potential that written work based on communicative tasks benefits from interaction and discussion about the content. A similar improvement was seen in the CRT group (RQ1b). In the case of the CRT condition the improvement was more significant, rising from an initial 50% for fluency to 63% at the delayed test. In comparison to the other groups which saw a rise of only 2-3%, this suggests that in answer to RQ1c, the interaction raised individual consciousness about errors which is then transferred to written outputs and an improvement in overall fluency and cohesion.

TABLE 8.
WRITTEN OUTPUTS BY GROUP

Group		Pre-test	Post-test	Delayed test
CRT	Fluency	50%	62%	63%
	Lexical richness	59%	62%	64%
	Accuracy	51%	62.5%	61.2%
CT	Fluency	70.1%	75%	73%
	Lexical richness	68.5%	85%	53%
	Accuracy	55%	56.1%	55.3%
CR & CT	Fluency	60%	62.1%	61%
	Lexical richness	55%	57%	43%
	Accuracy	61%	63.2%	79.4%
Control	Fluency	60%	66%	65%
	Lexical richness	60%	68%	53%
	Accuracy	55.6%	60%	48.75%

As with the speech fluency outcomes, and in regard to the research questions, the indication appears to be that all task types appear to deliver some improvement in written accuracy and fluency whilst the interaction continues, but that this may not be fully retained at the post-test stage. The loss seen in speech fluency however is less marked in the written outcomes, suggesting that improvements are maintained better with written than oral output, which is an area that could be investigated in future research. At the same time, examining the outcomes from the control group, there is again some natural learning but accuracy and lexical richness improvements following instruction were not maintained at the post-test stage. This again underlines the value of interaction and collaboration during task-based learning to embed the knowledge gained.

All groups showed improvement in fluency, but only the CRT improved in lexical richness, which suggests that raising consciousness improved access to new lexical items and variation following the interactions. It should also be noted that the written work in the combined group was marginally longer. Ellis (2009) suggests there is a relationship between complexity and fluency /accuracy which is reflected in the outcomes of this work. What is interesting in terms of written output is that the improvements were maintained at the delayed test stage, which suggests that the different processes required for written output, compared to oral may reinforce the improvements made as a result of interaction which is line with earlier work by Ellis and Barkhuizen, (2005).

E. Lexical Content of Speech Fluency

Given the aim of identifying whether interactional processes increased fluency and accuracy, the overall lexical content was evaluated, with the students divided into high, intermediate, and low proficiency groups. The statistical analyses consisted of a) correlations of groups and task performance measures to demonstrate the strength of the relationship between task type and various linguistic measures analysed in the three conditions and across modalities; b) analysis of variance with task type and modality as within-group and as between-group as a repeated measure to examine the impacts of task on learners' L2 production and c) one-way ANOVA tests to compare the performances of learners in the three conditions on the types of tasks and across modalities.

At the individual level, more advanced students appeared to engage in more pausing behaviour, which could be attributed to their longer utterance length, although this was not explicitly tested. Moreover, fluency is impacted on by personal speaking as well as language skills, and it should be recognised that in line with de Jong and Bosker (2013), hesitations and pauses are part of natural and successful communication approaches. With this in mind, the analysis presented here should be taken against this background understanding of the nature of fluent speech and how individual styles may vary.

In addition, there is a potential that the pausing of advanced students is an indication of other processes at work such as thinking and searching for a correct word, i.e. usage of interlanguage, or this may simply be a characteristic of their individual speech styles, which is a further factor in encouraging interaction and achieving positive outcomes. Again, this was not directly assessed so it would need further consideration in future studies. These tables, however, should be taken with caution due to the small size of each proficiency group. In other words, for this group at least the tables are indicative of the effectiveness of the interventions, and the valuable role of interaction in improving fluency.

Importantly the interactions created opportunity for learning, giving those of lower proficiency a learning space, where they could use their interactional skills and work together to improve fluency and accuracy, defined by Spitzberg and Cupach (2012) as interactional competence. What was also notable is some students were reticent about contributing to the interaction and only did so following encouragement from more proficient students, again supporting the role of interaction in improving fluency and accuracy, highlighting the value of mixed proficiency groups during TBL interventions. Furthermore, it should be note that there was higher proficiency and focused collaboration at the level of correction and confirmation. This was particularly noticeable within the CRT group who spent more time discussing potential solutions to the task than other groups, who had a more combative approach. This aligns with works relating to group familiarity and individual dynamics suggesting that the CRT group was more collaborative in their approach and reinforces the role of individual speaking styles and group dynamics.

F. Lexical Content of Written Fluency

According to Freed (2000) L2 fluency does not necessarily strictly develop in a linear manner and it refers to the final output rather than the actual writing process. For the written work, the examination was undertaken of the level of content words to identify the lexical density, can be shown in Table 9.

TABLE 9.
LEXICAL RICHNESS OF WRITTEN PRODUCTION

Group	Stage	Total Words	Total Content Words	Percentage of content words
CRT	Pre-Test	88	23	26%
CT	Pre-Test	68	20	29%
CRT & CT	Pre-Test	85	24	28%
CRT	Task	97	29	30%
CRT	Task	79	15	19%
CRT & CT	Task	95	22	23%
CRT & CT	Task	87	25	29%
CT	Task	84	17	20%
CT	Task	95	23	24%
CRT	Post-test	84	12	14%
CT	Post-test	108	29	27%
CRT & CT	Post-test	75	15	20%
CRT	Delayed Test	93	28	30%
CT	Delayed Test	98	28	29%
CT & CRT	Delayed Test	89	27	30%

*Note: In the task section, there are two scores as two tests were administered during the intervention

It is clear that the CRT group shows improvement from the pre-test during the task increasing percentage of content words from 26% to 30% of the overall text, based on type token ratio calculations but this was not reflected in the post-test stage when the percentage of content words had reduced to 21%. Interestingly there was improvement at the delayed test stage, for the CRT group returning to 30% by this stage, suggesting there was retention and learning that occurred. However, there is also the potential that following intervention the students had increased confidence in utilising more content words and awareness of the need to increase lexical richness in the compositions.

The CT group however showed no real change in their use of content words, suggesting that there was not the same level of improvement for this group. In the combined group however, there was a distinct improvement (of 2% from pre-test stage to delayed test – increasing from 28% to 30% during this period), again suggesting improved awareness of how to increase their lexical richness. However, the lexical richness was not supported by concurrent grammatical improvement across the groups. In addition, there was some improvement in English use, but the accuracy and proficiency of this longer term could not be assessed within the constraints of the current study.

Overall, no statistically significant differences were emerged between texts produced by the three conditions in terms of fluency and lexical richness. However, the texts produced by the combined group were more accurate than other groups. This suggests that task type may have an influence on the type and combination of strategies employed by the students. More specifically, those tasks that integrate spoken and written production may be more effective for L2 development than those that require only spoken production.

G. Accuracy during Interaction

Overall, the results reveal that the learners’ performance in the present study could not be accurate while they were pushed to produce the language using task-based sessions. This is in line with Foster & Skehan’s (1999) trade-off model although their model is focused on a general level of accuracy rather than at a more specific level. It was further assumed that there is no guarantee that an increase in fluency, lexical diversity, or even complexity would influence accuracy of any particular forms.

The results did however reveal that the students’ accuracy scores for L2 production seem to be closely related to how the learners are able to produce accurate sentences than the cognitive demands of task. In addition, it was identified that negotiation of meaning was not widely used, indeed, many of the students in the groups were reticent to indicate a lack of understanding, preferring to pretend that they understood. This could be an important factor in terms of managing interactions as it suggests that the personality differences and speaking styles of Saudi students are not always conducive to encouraging increased interaction and is something that teachers will need to take into account if their students are to benefit from the value of interaction during TBL interventions. Alongside the accuracy and lexical content evaluations, language episodes were also incorporated into the data collection and final analysis due to the indications from the literature review that they may influence how peers interact, specifically in regard to negotiation and achievement of accuracy (Watanabe and Swain, 2007).

H. Language-related Episodes (LREs)

Following Leeser (2004) and Swain & Lapkin (1998), the LREs may be identified as correctly resolved, unresolved, or incorrectly resolved. With this in mind, the identified LREs were coded based on these three potential outcomes and their linguistic focus. It is assumed that task type influenced the types of LREs as suggested by Kim (2008) and R ́ev ́esz (2011). So, the best way to represent the findings for LREs were to categorize them as grammar-focused (dealing with syntax or morphology), lexical-focused (dealing with word choices and word meaning), mechanics (dealing with spelling, pronunciation and punctuation), or pragmatic aspects. In regards to the research question, examination of the LREs can provide answer to whether fluency and accuracy improved in each of the group conditions. The results of LREs outcomes can be shown in Table 10.

TABLE 10.
LRES OUTCOMES

	CRT		CT		CRT & CT		Total
	N	%	N	%	N	%	
Number of times L1 used	56		93		93		242
Correctly resolved	40	61.54%	19	40.42%	38	50.7%	97
Unresolved/ abandoned	23	35.4%	22	46.8%	33	44%	78
Incorrectly resolved	2	3.08%	6	12.8%	4	5.3%	12
Total	65		47		75		187

The CRT students had the most correctly resolved LREs (n= 40, 61.54%) whilst the combined group had the highest number of unresolved episodes (n= 33, 44%) compared to the other groups. Part of the reason for this may have been the group dynamic, although there is also a potential that proficiency effects may have influenced these outcomes. It was also noted that there was a much higher use of L1 in both CT and combined groups (n= 93) compared to the CRT group (n= 56). This could be due to the task demands in the CT and combined students which required them to create new sentences and extend their L2 while the CRT students may refer to the sentences given in the worksheets. However, the use of L1 diminished over the intervention period in all groups, which are discussed later in this chapter suggesting

increasing confidence even when errors were made.

In terms of error type, the majority of the errors were grammatical in nature, such as omission of determiners or incorrect tense parsing. This appeared to be consistent across all conditions with no significant variation emerged between any of them. In addition, there was a high level of repetition in the interaction, again with no significant variations across the group. However, it is recognized that repeating phrases in examples such as 'go out, go out' 'where he, where he' is frequently used in natural speech to emphasize a point, or to confirm that what a speaker is saying is correct. In terms of new categories, which considered likely to occur prior to gathering the data, there were also indications that learners employed negotiation to determine the intended meaning, supporting again the role of interaction in improving accuracy and fluency.

In all three conditions, the most common LREs were grammatical in nature (n= 125), and predominantly focused on tense and pronoun errors. However, it was noted that in the CRT and combined groups, the number of correctly resolved events was significantly higher (n= 40 and 38 respectively) than that seen in the CT group (n=19). Furthermore, in the CRT and combined groups, there was a wider discussion on grammatical points (n= 92), suggesting a better focus on identifying errors and resolving them as a group. This suggests that there is a benefit to the CRT approach in achieving meaningful interaction and discussion leading to improvement in proficiency that is not so extensively present in the CT approach. This does not detract from the CT approach, but does underline the importance of development of the awareness raising of the students in identifying, discussing and correcting errors during task-based interaction.

Overall, in terms of LREs, discussions and interactions centred on the categories of incidence included meaning-based, grammatical, discourse, orthographic and identification of terms and individual constructs. The grammatical episodes were defined as the stages of the interaction where students engaged in discussion of morphological and syntactic elements of language. These instances were then subdivided into categories relating to: verb form (active/passive voice, aspect/tense forms, auxiliary verbs, and verb form), relative clauses (use of the defining/non-defining clause, choice of the relative pronoun/clause, omission of the pronoun, clause position: centre-embedded or right-embedded clauses, referring to the relative pronoun), subject-verb agreement, word order, choice of prepositions, conjunctions, definite or indefinite articles, gerunds or infinitive, and adverbs of time.

At the same time, in line with work by Benson, Pavitt & Jenkins (2005), the specific category was presented to enable coding of speech segments where the participants clearly identified phrases, words or sentences that were mentioned in task worksheet (CR tasks). Discourse episodes were considered to be those elements of the interaction where the students engaged in discussion of the older of sentences or sentence elements along with identifying parts that preceded or followed the parts under discussion. A few numbers of these instances were identified during the CR tasks, even though the majority of the activities did not require linking elements of text and discussions at the level of discourse which suggests that the interactions were focused on identifying errors and encouraging improvement in fluency and accuracy.

Overall, in terms of LREs (grammatical, lexical and pragmatic), it appears that the students were able to provide one another with mutual scaffolding during the interactions, which shows benefits of tasks and associated interactions can extend beyond grammar and lexis to L2 pronunciation or pragmatic features. The results also revealed that learners attended to lexical and grammatical items when they interacted with advanced interlocutors. Furthermore, it should be noted that mixed proficiency levels between the learners not only result in improvements in terms of attention to language but also to the nature of group work. The results lend support to Skehan's (1998) trade-off hypothesis. In other words, fluency and accuracy are dependent on task complexity and type, but can be supported by interaction and collaboration during the completion of tasks.

V. DISCUSSION

In this section, we discuss the findings of the study in relation to the research questions. Initially examining whether there were improvements in the fluency and accuracy of learner's L2 production, both oral and written, and then moving on to a discussion of negotiation of meaning. The results presented suggested two patterns of interest. The first concerns the comparisons seen between the three task conditions and any potential advantage of these, and the second is the potential influence of social dynamics and use of L1. Certainly the results indicate a value to the use of interaction to encourage negotiation of meaning and provision of learning spaces for mutual scaffolding during task-based activities.

The results of the pre-test/post-test/delayed test approach across grammatical knowledge, fluency and accuracy suggests that collaborative interactions do encourage a more fluent performance, but that once the interaction ceases, the improvement is not fully retained. In effect, and taking into account the outcomes for fluency and accuracy, it can be suggested that extending the task-based interaction, there would have been a concurrent on-going negotiation. The net result of this is that there could have been deeper learning and thus greater retention of the knowledge and greater attention to fluency and accuracy in both spoken and written outputs. This would however need to be reviewed in future research.

A factor indicated during evaluation of the interactions however is that the process of negotiation was not common, with students pretending to understand rather than request clarification. This could be due to individual personality

differences, cultural factors or potentially the setting, i.e. that of a secondary school, where students are not encouraged to engage in classroom debates and negotiate, and non-language lessons are more passive (Alharbi, 2015). Moreover, the secondary school setting for evaluation of the task-based approach and importantly interaction in the classroom has not been widely explored in the Saudi context and it is thus clear that further work is need (Al-Kathiri, 2015). Despite this overall reticence and lack of familiarity with collaborative working, it was noted that the group discussed features of the language during interactions, suggesting that with practice, the process would become more familiar and lead to increased negotiation.

Furthermore, in line with Spitzberg and Cupach (2012) it was identified that some learners found oral production challenging and daunting, reducing their motivation to engage fully in interactions, which had a negative impact on their overall fluency and accuracy. Again, this could be due to personality/learning styles as well as a lack of familiarity with a more student-centred approach and would need to be investigated in future research. However, it was further noted that in all three groups there were indications that the task type encouraged discussion of alternatives in terms of form and function and the introduction of options that could then be discussed and decided on. The value of this is that when groups are of mixed ability, as was the case for this study, new words and forms of expression were highlighted, clarified and learned by the lower-level learners, improving overall fluency and accuracy. For the higher proficiency learners, the interactions during the tasks supported an increase in their confidence about their own spoken fluency and accuracy, and provided them with an opportunity to self-reflect and evaluate, evidenced by pausing and other devices.

The increases in fluency and interaction over the intervention process further indicates that encouraging negotiation of linguistic features and their meaning in collaborative environments can encourage a deeper understanding that consolidates the cognitive processing necessary for effective retention and subsequent use of the knowledge. In this respect, what is particularly critical is that when learners produce language in collaborative work conditions, they do not just produce output, but gain a benefit from the monitoring and feedback they receive from interlocutors whilst attempting to produce accurate verbal utterances or written work. It appears therefore that in the collaborative setting, there is a greater individual focus on fluency from the students, which is important in terms of teaching practice. In other words, in line with Tavakoli and Hunter (2018), there needs to be an increased focus of the achievement of fluency and accuracy in the development of classroom instruction and tasks given.

The observations of the students in this study are thus in line with Swain and Lapkin (1998) and Swain and Watanabe (2012) who noted that during interaction learners pay attention to the forms used by others in their group, develop the ability to evaluate accuracy and appropriateness and in so doing can provide feedback and correction. In effect, there are two monitoring levels, self and others, which occur simultaneously, embedding information in the language centres of the brain and promoting more accurate, high quality task outcomes.

In addition, the current study confirms the work of Witton-Davies (2014) who identified that dialogue interactions are more fluent than monologues, including reduced pausing, elevated speech rates and a reduction in repair words. However, this is dependent on the task type and individual differences as well as the linguistic features. Although the students in this study did not, as such, engage in monologues, it can be seen that this is confirmed when some individual students engaged in longer utterances. Moreover, whilst turn-taking patterns and interruptions were not directly analysed for this work, several instances of students indicating '*it's my turn*' or similar were identified during the interaction. This suggests that across all groups there was recognition of the need for dialogue and collaborative interaction to achieve effective task outcomes (Cameron, 2001).

Although the quantitative analysis suggested that overall interaction improved fluency and accuracy, it should also be recognised that the dominance and passivity of individual group members may vary, impacting on overall turn-taking patterns. Although this element was not statistically measured, it appeared to conform to the work of Hudson-Kam and Edwards (2008) regarding balance within interactions. Overall however it can be indicated that the majority of students felt relatively confident in their ability to promote fluency in class, although it is also recognised that many of the proposed activities were not directly tailored to fluency and accuracy and as such were more suitable for the development of a more general speaking ability. Furthermore, in the current study, there were cases identified from the transcripts where some students appeared to dominate, or had to encourage participation from others. In terms of fluency measurement therefore, this study could support previous findings regarding turn-taking and interruptions.

VI. IMPLICATION

This work has identified some clear benefits to the use of CT and CRT tasks in promoting interaction and thus development of fluency and accuracy. What this suggests and what we therefore propose is that Saudi secondary schools should place greater focus on developing a more task-based interaction focused curriculum to encourage discussion between students that leads to improvements through negotiation, self-repair and correction and ultimately improvements in their fluency and accuracy. At the same time, there is an implication for curriculum and the importance of ensuring positive group dynamics, balancing proficient, intermediate and low level students to encourage development of learning.

VII. CONCLUSION

Task-based learning as a beneficial means of improving second language knowledge is not in dispute, nor is the value of including either consciousness raising or communicative tasks within the process. What has been determined from this work is that the interaction that occurs in the completion of these tasks is a potential factor in increasing fluency and accuracy, through the process of negotiation of meaning and overall increasing of ability and confidence amongst students. At the same time, there is also a need for additional work to further consider the need for use of L1 and the wider impact of factors such as task type, group dynamics and how much impact language related episodes can have on developing fluency and accuracy.

Although there is definitely value in the work undertaken and an indication of potential value within the Saudi school system, we do recognise that the work has limitations. Firstly, this was a small scale study over a defined period of time and further research is likely to be necessary to verify the findings and examine factors such as use of L1 and the assessment of LREs as a measure of fluency before concrete steps can be taken to change the curriculum directly. This having been said, we believe that the findings present some interesting discussion points for the writers of secondary school EFL curricula in Saudi Arabia.

REFERENCES

- [1] Alharbi, H. A. (2015). Improving Students' English Speaking Proficiency in Saudi Public Schools. *International Journal of Instruction*, 8.1, 105-116.
- [2] Al-Kathiri, F. (2015). Beyond the classroom walls: Edmodo in Saudi secondary school EFL instruction, attitudes and challenges. *English Language Teaching*, 8.1, 189-204
- [3] Belbin, R. M. (2010). Team roles at work. Oxon: Routledge.
- [4] Block, D. (2003). The social turn in second language acquisition. Edinburgh: Edinburgh University Press.
- [5] Cameron, L. (2001). Teaching languages to young learners. Cambridge: Cambridge University Press
- [6] De Jong, H., & Bosker, R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In Eklund R. (ed.) *Proceedings of Disfluency in Spontaneous Speech Diss 2013* Stockholm: Royal Institute of Technology (KTH). 17-20.
- [7] De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36.2, 223-243.
- [8] De la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition*, 24.1, 81-112.
- [9] Elliott, R. W., & Zhang, D. (2019). Cognitive perceptions of second language acquisition in Chinese undergraduate ESL/EFL students: A mixed methods approach. *Int J Eval & Res Educ*.
- [10] Ellis, R. (1991). Grammar teaching practice or consciousness-raising? In R. Ellis (ed.), *Second Language Acquisition and Second Language Pedagogy*. Clevedon, Avon: Multilingual Matters, 232-241.
- [11] Ellis, R. (2009). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19.3, 221-243.
- [12] Ellis, R., & Barkhuizen, G. (2005). Analysing learner language. Oxford: Oxford University Press.
- [13] Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21.2, 285-301.
- [14] Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3, pp. 215-247.
- [15] Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98-116.
- [16] Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28.2, 323-351.
- [17] Gass, S., Mackey, A., & Ross-Feldman, L. (2011). Task-Based Interactions in Classroom and Laboratory Settings. *Language Learning*, 61(supp. 1), 189-220.
- [18] Gorse, C. A., & Sanderson, A. M. (2007). Exploring group work dynamics. Paper presented at the In Proceedings 23rd annual ARCOM conference, ed. D. Boyd, 295-304, Belfast, UK.
- [19] Han, H. (2014). Transforming EFL Classes from Lecturing to Cooperative Learning. *Journal of Language Teaching & Research*, 5.4, 948-952.
- [20] Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, 2.1, 1-23.
- [21] Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40.1, 109-131.
- [22] Hmelo-Silver, C., Chernobilsky, E., & Jordan, R. (2008). Understanding collaborative learning processes in new learning environments. *Instructional Science*, 36.5, 409-430.
- [23] Huba, M. E., & Freed, J. E. (2000). Learner-centered assessment on college campuses: Shifting the focus from teaching to learning. Allyn & Bacon, 160 Gould St., Needham Heights, MA 02494.
- [24] Hudson-Kam, C., & Edwards, N. (2008). The use of uh and um by 3-and 4-year-old native English-speaking children: Not quite right but not completely wrong. *First language*, 28.3, 313-327.
- [25] Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24. 4, 541-577.
- [26] Jeon, I., & Hahn, J. (2006). Exploring EFL teachers' perceptions of task-based language teaching: A Case study of Korean secondary school classroom practice. *Asian EFL Journal*, 8.1, 123-139.
- [27] Jong, Y.O., (2009). An investigation into the benefits of collaborative writing for the development of EFL children's communication skills: A reflective report of a teacher researcher (Doctoral dissertation, University of Warwick).
- [28] Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern*

- Language Journal*, 92.1, 14-130.
- [29] Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3.2, 73-93.
- [30] Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8.1, 55-81.
- [31] Long, M. H. (Ed.) (1996). *The role of the linguistic environment in second language acquisition* (Vol. 2). New York: Academic Press.
- [32] Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96.2, 190-208.
- [33] McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32.2, 207-224.
- [34] Mohamed, N. (2004). Consciousness-raising tasks: a learner perspective. *English Language Teaching Journal*, 58.3, 228-237.
- [35] Nassaji, H., & Tian, J. (2010). Collaborative and individual output task and their effects on learning English phrasal verbs. *Language Teaching Research*, 14.4, 397-420.
- [36] Nassaji, H., & Tian, J. (2014). The Role of Language Coproduction in Learning English Vocabulary *Procedia-Social and Behavioral Sciences*, 143, 794-798.
- [37] Puntambekar, S. (2006). Analyzing collaborative interactions: divergence, shared understanding and construction of knowledge. *Computers & Education*, 47.3, 332-351.
- [38] Révész, A. (2011). Task Complexity, Focus on L2 Constructions, and Individual Differences: A Classroom-Based Study. *Modern Language Journal*, 95 (SUPPL. 1), 162-181
- [39] Saito, K., Trofimovich, P. & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37.2, 217-240.
- [40] Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20.4, 286-305.
- [41] Spada, N., & Lightbown, P.M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, 15.02, 205-224.
- [42] Spitzberg, B., & Cupach, W. (2012). *Handbook of interpersonal competence research*. Berlin: Springer Science & Business Media.
- [43] Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching*, 46.2, 195-207.
- [44] Swain, M., & Lapkin, S. (1998). Interaction and second language learning: two adolescent French immersion students working together. *Modern Language Journal*, 82, 320-337.
- [45] Swain, M., & Watanabe, Y. (Eds.). (2012). *Language: Collaborative Dialogue as a Source of Second Language Learning*. Hoboken, NJ: Wiley-Blackwell.
- [46] Trofimovich, P., McDonough, K., Dao, P., & Abashidze, D. (2019). Attitudinal bias, individual differences, and second language speakers' interactional performance. *Applied Linguistics Review*. Published online 14 June 2019. <https://doi.org/10.1515/applirev-2019-0010>.
- [47] Van Lier, L. (Ed.) (2001). *Language awareness*. Cambridge: Cambridge University Press.
- [48] Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11.2, 212-142.
- [49] Witton-Davies, G. (2014). *The study of fluency and its development in monologue and dialogue*. Lancaster University. Lancaster, UK.
- [50] Wray, A. (2002). *Formulaic language and the lexicon* (V.3). Cambridge, UK: Cambridge University Press.
- [51] Zastrow, C. (2009). *Social work with groups: A comprehensive workbook 7th (ed.)*. Belmont, CA: Brooks/Cole.

Adnan Mukhrib is an assistant professor of Applied Linguistics and TESOL at Umm AL Qur'a University, English language centre, Makkah, Saudi Arabia. He completed a PhD in the areas of second language acquisition from the University of Bristol, UK, did an MA in Applied Linguistics and TESOL at Newcastle University, and his BA degree in English language and Linguistics from Umm AL- Qur'a University, Makkah, Saudi Arabia. Prior to joining the English language centre at Umm Al Qur'a University, he taught a variety of courses at Yanbu University College, including phonetics & phonology, Introduction to Applied Linguistics, sociolinguistics, pragmatics, Linguistics and academic writing in Applied Linguistics department at Yanbu University College. More recently he worked as an EAP tutor at the University of Nottingham, and Southampton University in the UK, and also taught academic writing for undergraduate programs in Yanbu Industrial College, Yanbu, Saudi Arabia.

His research interests lie in the areas of second language acquisition. In particular, he conducts research on the roles of tasks, input, interaction and individual differences in L2 development. He is also interested in aspects of second language acquisition including fluency, pronunciation, accuracy and complexity of L2 speech production and comprehension, the cognitive processes underlying L2 learning. His current projects investigate the cognitive processes involved in second language writing and speaking, the role of output-based tasks in second language learning, and sources of difficulty in second language production.