Research Article 10.30495/JAL.2020.683555

The Relationship Between Three Components of CAF in Learners' Oral Performance at Different Proficiency Levels

Kazem Najmi [1], Masood Siyyari [2*], Gholam-Reza Abassian [3]

[1, 2, 3] *Department of English, Science and Research Branch, Islamic Azad University, Tehran, Iran*
[3]*Department of English, Imam Ali University, Tehran, Iran*
*Corresponding author: siyyari@gmail.com
(Received: 2021/2/15; Accepted: 2021/7/5)

Online publication: 2021/7/31

Abstract

The current study aimed at finding out the relationship between complexity, accuracy, and fluency (CAF) in speaking performance across language proficiency levels with the focus on Skehan's Trade-off Hypothesis. The selected participants of the present study included 60 language learners who were selected out of 90 learners. Through the results obtained from Oxford Placement Test, 21 participants were placed at elementary, 20 at intermediate, and 19 at advanced levels. The nationality of all participants was Iranian and therefore they shared similar L1 background. The sample participants were asked to perform the designed speaking tasks in different task complexities (low, mid, and high) and their oral CAF were measured and analyzed. To measure CAF, the percentage of error free C-units for accuracy, clauses per C-units for grammatical complexity, type-token ratio (TTR) for lexical complexity, and total number of tokens (words)/ total task time (per minutes) for fluency were used. The results of Pearson product-moment correlation revealed a significant positive relationship among complexity, accuracy, and fluency in oral performance. The results of the study indicate that there is variation in the process of gradual complexification in spoken L2 production across proficiency levels. In fact, the results of this study revealed distinguishing features in all three CAF components. Moreover, the findings of the present study provide pedagogical implications and recommendations for teachers, syllabus designers, and language assessors. Finally, some suggestions for other interested researchers in this field are presented.

Keywords: accuracy, complexity, fluency, task complexity, trade-off hypothesis

Introduction

Many researchers accept limitations in learner language performance. Simply put, emphasizing one element of language performance may result in a lower performance in one or both of the other components. From a cognitive, information processing framework, Skehan (2009) believes that there is a competitive relationship between Complexity, Accuracy, and Fluency (CAF) components due to limited mental resources, specifically limited attentional capacity and working memory. In Skehan's limited attentional model, this limited capacity in online processing is because of a single-source view of attention. If trade-off effects are expected, a theory ought to predict which CAF constructs are likely to show the effects and why. Skehan (1998b) mentions that adult learners emphasize meaning over form, which can potentially hinder language development. Then, when students focus on form, there is a secondary contrast between control of form (accuracy) and inter-language risk-taking (complexity). All language learners have these tensions during performance. When a performance indicates improvement, rather than trade-off effects, in two areas, according to Skehan (2009a) two explanations are possible. The growth in two areas could actually be the result of separate influences. For example, the task structure may help accuracy when the information manipulation during the task needs the learners to utilize subordination that increases grammatical complexity. Alternatively, while analyzing group data, some learners might focus on one area of the CAF triad while others attend to another area. Collected data may then provide the appearance that two areas, which should be in a competitive relationship, are both showing improvement. Consequently, correlations must be run on learners' performances, not just at group levels, he adds.

Skehan (2009b) detects that human working memory and attentional capacity are limited and that when learners direct their attention to one dimension of CAF, it might reduce their attention for other areas—the Trade-off Hypothesis (Skehan, 2009b). He also mentions that 'learners can prioritize attention to particular areas' and that '. . . task performance appears to be the result of an interpretation by the task participants should do' (p. 115). In other words, it is the learners who change their goals and

prioritize certain areas, complexity or accuracy, and that what they choose to prioritize is different for different learners; some prioritize complexity and some accuracy. Skehan (2014) believes that different participants seem to focus on different aspects of form. The trade-off effects in complex tasks, such as online oral language performance, are found in many studies (Ellis & Barkhuizen, 2005; Skehan, 2015; Skehan & Foster, 1997) researching CAF from instructed language learning settings. In a psycholinguistic view of language proficiency, these research results are referred to the learners' inability to attend to all components of language performance simultaneously at the highest possible level. Thus, learners must prioritize one component of the language performance. However, the field has not reached a conclusion regarding what components really trade-off because of differences in task complexity and task conditions (Ellis & Barkhuizen, 2005).

Many researchers believe in limitations in learners' oral performance. In fact, focusing on one component of language performance may result in a lower performance in one or both of the other components. Skehan (2009a) portends a competitive relationship between CAF because of mental resources limitations, specially limited attentional capacity. If trade-off effects are expected, a theory should predict which CAF constructs are likely to indicate the consequences and why. While analyzing group data, some learners might attend to one particular area of the CAF triad, whereas others attend to another area. Collected data might then give the appearance that two areas, which should be in a competitive relationship, are both showing improvement. Even for researchers who do not believe a single-source capacity limitation, trade-off effects might be seen in language performance, but these trade-off effects can be clarified by attention.

After measurements for complexity, accuracy, and fluency are selected, researches (Kim, Nam, & Lee, 2016; Skehan, 2015) have usually thought about the way these constructs of language performance interact. This provides a review of researches which looked at potential trade-off effects (where a higher performance in one component corresponds to a lower performance in another) between complexity, accuracy, and fluency during language performance. The review focuses on research of oral language performance, and particularly in learners of English.

Robinson's cognition hypothesis (Robinson & Gilabert, 2007) states that increased accuracy and complexity will be encouraged by increasing the cognitive demands of the task provided to individuals. As the learners try to produce the language needed by the greater functional demands in the relatively increased complexity of the task, their language performance will improve.

It is accepted in the field that learners cannot attend to all areas of CAF performance, especially in demanding tasks (Robinson, 2011; Robinson, Cadierno, & Shirai, 2009). Although the findings are different and sometimes there are some contradictions, some trade-off effects in language performance are anticipated. Many studies (Schmidt, 2001; Skehan, 2015) with cross-sectional research designs report a trade- off between accuracy and fluency, while Mizera's (2006) findings, based on correlations of learners' performance, recommend that these are connected growers. Moreover, studies with between-group designs (Ahmadian & Tavakoli, 2011; Yuan & Ellis, 2003) have found that individuals can have higher accuracy and complexity at the expense of fluency. However, an emerging key explanatory variable is the task given to the student and the research design.

The construct of L2 performance and proficiency has long been recognized as multi-componential and multi-dimensional, comprising three principal components: complexity, accuracy, and fluency (Skehan, 1998; Ellis, 2008; Ellis & Barkhuizen, 2005). As such, CAF features have been widely used to characterize test performances and test-taker proficiency levels in both L2 speaking and writing assessments (Housen & Kuiken, 2009).

Trade-off Effects Predicted in Language Performance

Skehan (1998a) emphasizes that adult students are various in their learning style by learning through exemplars and putting emphasis on fluency or by learning through analysis and emphasizing complexity or accuracy. This meaning versus form dichotomy has also been studied as a limitation in attending to information (e.g., VanPatten's input processing theory). Pedagogy also echoes the fluency-form distinction, in which spontaneous, free-flowing language is the goal of fluency-oriented tasks and

a focus on form and control is the goal of accuracy-oriented tasks (Ruiz-Funes, 2015). Empirical findings have supported the form-meaning dichotomy. In a study looking at the effect of task repetition, Bygate (1999) found that grammatical complexity increased but at the expense of fluency, measured by the number of pauses (Brumfit, 1984; Bygate, 1999; VanPatten, 2012).

As it was mentioned earlier, accuracy and complexity may compete during oral language performance. Skehan and Foster (1997) reported a trade-off between accuracy and complexity in a study looking at the effect of planning during three oral tasks. The means of planning group, the group who had some minutes to plan their talk, on all measures were higher than the non-planning group, i.e. those who did not have the chance to plan their talk. In the decision-making tasks, the planning group significantly outperformed the non-planning group on the complexity measure but not the accuracy measure while on the narrative task; the planning group significantly outperformed the non-planning group in accuracy but not in complexity. Also, during the personal information task, the planning group significantly outperformed the non-planning group on all three measures.

CAF and the Level of Proficiency

The interest in researching second language oral CAF has increased over the past decades because of the significant role it plays in reflecting the development of communicative ability and assessing learner proficiency. From a developmental perspective, when L2 learners expand their L2 repertoire and progress to a higher proficiency level, their language use becomes more automatic and they typically produce output of higher fluency, accuracy, and complexity (Housen, Kuiken, & Vedder, 2012; Skehan, 2009, 2015). The concept of CAF is therefore related to proficiency level where a positive correlation is usually assumed between the two. In L2 assessment, CAF has long been recognized as a key construct that reflects L2 proficiency (see Fulcher, 2003). Researchers commonly agree that CAF is a complex and multifaceted construct, often difficult to define and measure (Kormos, & Dénes 2006; Lennon, 1990; Segalowitz, 2017). In recent years, however, attempts have been made to unpack the concept and identify ways of measuring it reliably. Segalowitz's (2010) model of CAF and Skehan's (2003) framework for finding the relationship between the

three components of CAF are two examples of successful attempts that have expanded our conceptual understanding of CAF.

More recently, other researchers (Bosker et al., 2013; Huensch & Tracy–Ventura, 2016; Hunter, 2017; Skehan, 2015) have argued that a conceptualizing and measuring CAF at a fine-grained level can not only reveal more about the connection between L2 speech and the underlying speech production processes (Huensch & Tracy–Ventura, 2017; Hunter, 2017; Tavakoli & Hunter, 2018), but it will also enhance a more reliable understanding of what characterizes CAF at different levels of proficiency, making speaking-test rating scales more useful and meaningful for users and examiners (Nakatsuhara, 2014; Tavakoli, Nakatsuhara, & Hunter, 2017).

Still a crucial question remains unanswered here: what is the relationship between components of CAF in learners with different level of proficiency (e.g., elementary, intermediate, and upper-intermediate)?

The present study aimed to find out if Skehan's Trade-off hypothesis existed between complexity, accuracy, and fluency. Moreover, Skehan (2009) considered complexity as a general component, whereas in the present study in order to have a deep understanding of the issue, complexity was separated as lexical and grammatical complexity. Learners' initial proficiency level was another focus of the present study. Hence, the following research questions (RQ) were proposed.

RQ_1: Is there a relationship between complexity and accuracy in elementary, intermediate, and advanced EFL learners' oral performance?

RQ_2: Is there a relationship between complexity and fluency in elementary, intermediate, and advanced EFL learners' oral performance?

RQ_3: Is there a relationship between fluency and accuracy in elementary, intermediate, and advanced EFL learners' oral performance?

And the following null hypotheses (H0) were introduced.

$H0_1$: There is not any relationship between complexity and accuracy in elementary, intermediate, and advanced EFL learners' oral performance.

$H0_2$: There is not any relationship between complexity and fluency in elementary, intermediate, and advanced EFL learners' oral performance.

$H0_3$: There is not any relationship between fluency and accuracy in elementary, intermediate, and advanced EFL learners' oral performance.

Method

Participants

A total number of 60 EFL learners from Afaq Language Center located in Tehran constituted the participants of the study. To increase the practicality and manageability of the research, all the participants were randomly selected from the mentioned institute. All the participants study English solely within the educational system of Iran and none of them had the experience of studying or living in English-speaking countries. Participants were between 18 to 22 years old. The participants' English proficiency was estimated to be at three levels of elementary, intermediate, and advanced based on the results of Oxford Placement Test (2004) placement test.

Instrumentations

Oxford Placement Tests

Allen's (2004) Oxford Placement Test (OPT) provides researchers with a reliable and efficient means of placing learners at the beginning of a research course. The test includes 200 grammar and listening questions. OPT was administered to 60 learners to place them in three proficiency levels, 21 learners in elementary level, 20 in inter-mediate level, and 19 in advanced level. The criterion of this classification was SD which is shown in Tables 2 and 3.

Speaking Tasks

In order to design suitable speaking tasks related to the learners' current level of proficiency, available books in the market were used. In such books, there are different speaking tasks. Since these books, namely, American English Files, English Results, Top Notches, and York Mission Possible books, are at different levels, it was more precise to choose the speaking tasks (i.e. personal information exchange, narrative, and decision-making) of these books based on the learners' current level of proficiency, that is, levels A, B and C.

Procedure

The Oxford Placement Tests (OPTs) provide teachers with a reliable and efficient means of placing learners at the start of a course of a research. The tests have been calibrated against the levels system provided by the Common European Framework of Reference for Language: Learning, Teaching, Assessment (commonly known as the CEF), which has been

adopted by the Association of Language Testers in Europe (ALTE) as well as by government and major institutions, including exam boards, throughout Europe. The OPTs can clearly and reliably identify any learner's CEF level (on the A1 to C2 CEF scale) and also provide a score which will show where the learner is within that band. The test includes 200 questions which are divided into listening test and grammar test with 100 questions in each. Both parts were administered prior to the course to guarantee the participants' homogeneity in terms of English proficiency level. OPT was administered to 60 participants to place them in the three proficiency levels, that is, elementary, intermediate, and advanced. Based on the obtained scores, learners were placed at three levels: 21 learners at elementary level, 20 learners at intermediate level and 19 learners at advanced level. In order to increase the manageability of the study, instead of placing the learners into six levels of A1, A2, B1, B2, C1, and C2 based on their test scores, it was decided to merge A1 and A2 together. The same decision was also made about B and C. Therefore, level A, elementary, were those learners whose scores fell between 90 to 119, level B, inter-mediate, were those learners whose scores fell between 120 to 149 and level C, advanced, were those learners whose scores fell between 150 to189.

To answer the research questions, the correlation between the three components of CAF at all three levels was measured.

Accuracy Measures

To measure accuracy in this study, the percentage of the error free C-units (Robinson, 2003) was used.

Complexity Measures

To measure grammatical complexity, the clauses per C-unit and to measure lexical complexity, type-token ratio (Robinson, 2003) were used.

Fluency Measures

To measure fluency, the number of tokens (words) was divided by the total task time in minutes (Robinson, 2003).

Statistical Analysis

In order to address the research questions, Pearson product-moment correlation coefficient was calculated. It should be mentioned that the correlation between three components of CAF and the three levels were not

reported separately. There were three levels just to make sure that learners were doing the tasks related to their current levels of proficiency.

Results

OPT Homogeneity Test Results

OPT was administered to 90 participants to place them in the three proficiency levels, that is, elementary, intermediate, and advanced levels. The descriptive statistics in Table1 indicates that the mean, median and mode of the OPT scores were 123.26, 119, and 89 respectively.

Table 1
Descriptive Statistics for OPT Score

N	Mean	Median	Mode	SD	Kolmogorov-Smirnov Z	Sig. (2-tailed)
60	123.26	119	89	37.04	.989	.282

Moreover, Table 1 reflects that the normality of the scores was at a significance level for Kolmogorov-Smirnov Z test of normality (p = .28) which was greater than .05 and not significant.

Table 2 illustrates the descriptive statistics for the OPT scores across three proficiency levels.

Table 2
Descriptive Statistics for OPT Score across Three Proficiency Levels

Level	N	Range score	Mean	SD	SEM
Elementary	21	90 – 119	103.38	8.506	1.856
Intermediate	20	120 – 149	133.30	9.761	2.183
Advanced	19	150 – 189	166.37	10.745	2.465

As seen in Table 2, among 60 students, based on the norm suggested by OPT, those students whose OPT scores were between 90 and 119 were chosen as elementary participants ($\bar{x}$ = 103.38, SD = 8.51, n = 21), those students who scored between 120 and 149 were selected as intermediate ($\bar{x}$ = 133.30, SD = 9.76, n = 20), and those learners whose scores were

between 150 to 189 were considered as advanced participants ($\bar{x}$ = 166.37, *SD* = 10.74, *n* = 19).

The Relationship between Complexity and Accuracy in Oral Performance

The first research question of this study aimed at investigating whether there is any significant relationship between complexity and accuracy in the oral performance. Pearson product-moment correlation coefficient was performed to answer this research question. As Field (2009) believes, before one decides to apply parametric tests, three assumptions (interval data, independence of subjects, and normality) should not be violated. The first assumption is met because the present data are measured on an interval scale (Pallant, 2013).

Also Bachman (2004) notifies that the assumption of independence of subjects is met when "the performance of any given individual is independent of the performance of other individuals" (p. 236). The third assumption deals with the normality of the data which is tested via One-Sample Kolmogorov-Smirnov Test. Table 3 manifests the results of the normality test for the two main variables of the present study (Bachman, 2004).

Table 3
One-Sample Kolmogorov-Smirnov Test for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Accuracy Scores

Variable	*N*	Kolmogorov-Smirnov Z	Sig.
Oral grammatical complexity	60	.808	.530
Oral lexical complexity	60	.510	.894
Oral accuracy	60	.569	.853

As illustrated in Table 3, One-Sample Kolmogorov-Smirnov Test revealed that the scores of the three variables, that is, oral grammatical complexity (p = .53, p > .05), oral lexical complexity (p = .89, p > .05), and oral accuracy (p = .85, p > .05) satisfy normality assumptions. Therefore, the parametric Pearson Product-Moment Correlation Coefficient was used to measure the relationship between the three main variables of the study.

Before explaining the results of Pearson Product-Moment Correlation, the descriptive statistics for the scores for oral grammatical and lexical complexity and oral accuracy were computed and given in Table 4.

Table 4
Descriptive Statistics for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Accuracy Scores

Variable	*N*	Mean	SD	Std. Error
Oral grammatical complexity	60	2.566	.549	.070
Oral lexical complexity	60	66.608	9.037	1.166
Oral accuracy	60	59.568	5.798	.748

Table 4 manifests the mean and standard deviation for the oral grammatical complexity scores ($\bar{x}$ = 2.57, *SD* = .55, *n* = 60), for the oral lexical complexity scores ($\bar{x}$ = 66.61, *SD* = 9.04, *n* = 60), and for the oral accuracy scores ($\bar{x}$ = 59.57, *SD* = 5.80, *n* = 60). (See Appendix F for the raw scores gained on oral grammatical and lexical complexity and oral accuracy).

As it is evident in the scatter plot (Figure 1), we can draw a straight line through the main cluster of the points signifying a linear relationship. Thus, the linearity assumption of the two pairs is met for performing Pearson product-moment correlation coefficient. Furthermore, the shapes of the cluster are even from one end to the other in the scatter plot. Consequently, our data met the homoscedasticity assumption.

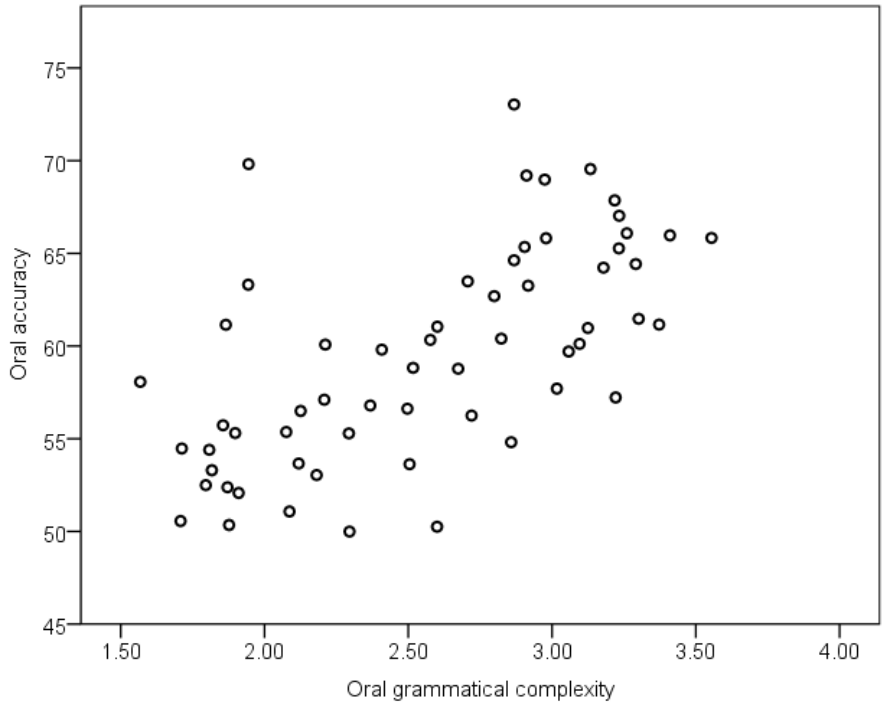

Figure 1. Scatter Plot of the Correlation between Oral Grammatical Complexity and Oral Accuracy

Additionally, the scatter plot (Figure 1) can tell us if the relationship between the two variables is positive or negative. In the scatter plot below, the direction of the line that is drawn through the points, point from left to right upward in the scatter plot. This left to right upward trend reflects a positive relationship; high scores on X, oral grammatical complexity correlates with high scores on Y, oral accuracy.

Moreover, the scatter plot of the correlation between oral lexical complexity and accuracy (Figure 2) was made before running correlation.

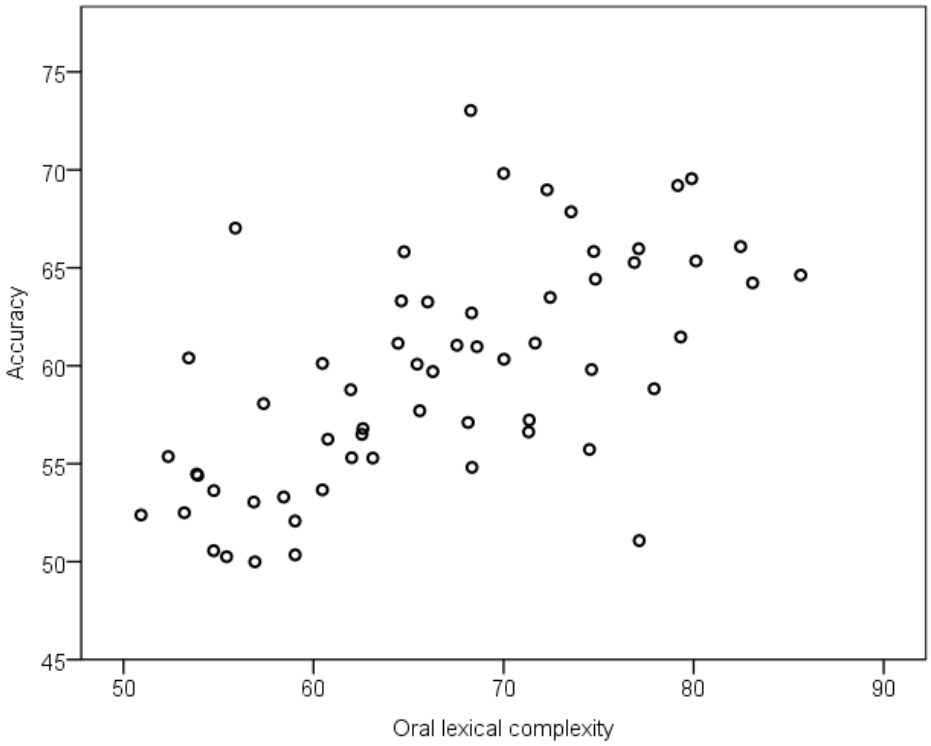

Figure 2. Scatter Plot of the Correlation between Oral Lexical Complexity and Oral Accuracy

As it is shown in the scatter plots, we can draw a straight line through the main cluster of the points showing a linear relationship. Therefore, the linearity assumption of the two pairs is not violated for running Pearson product-moment correlation coefficient. Furthermore, the shapes of the cluster are even from one end to the other in the scatter plot; so, our data enjoy the homoscedasticity assumption as well. Besides, in the scatter plot (Figure 2), the direction of the line that is drawn through the points, point from left to right upward in the scatter plot denotes a positive relationship: high scores on X, oral lexical complexity correlates with high scores on Y, oral accuracy.

Table 5 represents the results of Pearson product-moment correlation between the three variables. As it is evident from Table 5, Pearson correlation found a significant positive relationship between oral grammatical complexity and oral accuracy (r = .64, n = 60) at the significance level of .000 < .05, with high levels of oral grammatical

complexity associated with high levels of oral accuracy. Also *r* squared was computed as .40 using standard regression indicating that 40% of the variance of oral accuracy is accounted for or predicted by grammatical complexity. This effect size is medium based on Cohen's guideline (1998, pp.79-81). In fact, according to Cohen, the effect sizes for Pearson correlation that lie between .10 and .29 are considered small, between .30 and .49 they are seen medium, and finally between .50 and 1.0 they are called large effect size.

Table 5
Pearson Correlation for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Accuracy Scores

		Oral accuracy
Oral grammatical complexity	Pearson Correlation	.637**
	Sig. (2-tailed)	.000
	N	60
Oral lexical complexity	Pearson Correlation	.622**
	Sig. (2-tailed)	.000
	N	60

In addition, as seen in Table 5, a significant positive relationship was detected between oral lexical complexity and oral accuracy (r = .62, n = 60) at the significance level of .000 < .05, with high levels of oral lexical complexity associated with high levels of oral accuracy. Moreover, *r* squared was calculated as .39 through standard regression showing that 39% of the variance of oral accuracy is predicted by lexical complexity. This effect size is medium based on Cohen's guideline (1998, pp.79-81).

The Relationship between Complexity and Fluency in Oral Performance

The second research question of this study aimed at investigating whether there is any significant relationship between complexity and fluency in oral performance. Pearson product-moment correlation coefficient was performed to answer this research question. As it is obvious in Table 6, One-Sample Kolmogorov-Smirnov Test revealed that the scores of the three variables, that is, oral grammatical complexity (p = .53, p > .05), oral lexical complexity (p = .89, p > .05), and oral fluency (p = .90, p > .05) satisfy

normality assumptions. Therefore, the parametric Pearson Product-Moment Correlation Coefficient was used to measure the relationship between the three main variables of the study.

Table 6
One-Sample Kolmogorov-Smirnov Test for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Fluency Scores

Variable	*N*	Kolmogorov-Smirnov Z	Sig.
Oral grammatical complexity	60	.808	.530
Oral lexical complexity	60	.510	.894
Oral fluency	60	.500	.903

Before explaining the results of Pearson Product-Moment Correlation, the descriptive statistics for the scores of oral grammatical and lexical complexity and oral fluency were computed and given in Table 7.

Table 7
Descriptive Statistics for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Fluency Scores

Variable	*N*	Mean	SD	Std. Error
Oral grammatical complexity	60	2.566	.549	.070
Oral lexical complexity	60	66.608	9.037	1.166
Oral fluency	60	77.412	9.051	1.168

Table 7 indicates the mean and standard deviation for the oral grammatical complexity scores ($\bar{x}$ = 2.57, *SD* = .55, *n* = 60), for the oral lexical complexity scores ($\bar{x}$ = 66.61, *SD* = 9.04, *n* = 60), and for the oral fluency scores ($\bar{x}$ = 77.41, *SD* = 9.05, *n* = 60). (See Appendix F for the raw scores gained on oral grammatical and lexical complexity and oral fluency).

The scatter plot of the correlation between oral grammatical complexity and fluency (Figure 3) was created before running correlation. As it is evident in the scatter plot, we can draw a straight line through the main cluster of the points signifying a linear relationship. Thus, the linearity assumption of the two pairs is met for performing Pearson product-moment correlation coefficient. Furthermore, the shapes of the cluster are even from

one end to the other in the scatter plot. Consequently, our data met the homoscedasticity assumption.

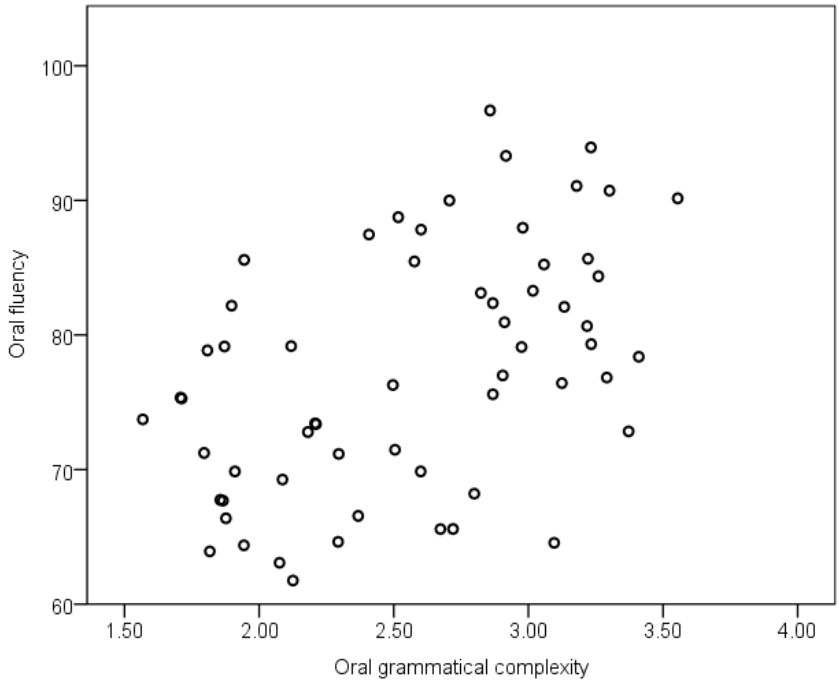

Figure 3. Scatter plot of the correlation between oral grammatical complexity and oral fluency

In addition, the scatter plot (Figure 3) can tell us if the relationship between the two variables is positive or negative. As represented in Figure 3, the direction of the line that is drawn through the points, is from left to right upward in the scatter plot. This left to right upward trend reflects a positive relationship; high scores on X, oral grammatical complexity correlates with high scores on Y, oral fluency. Moreover, the scatter plot of the correlation between oral lexical complexity and fluency (Figure 4) was made before running correlation.

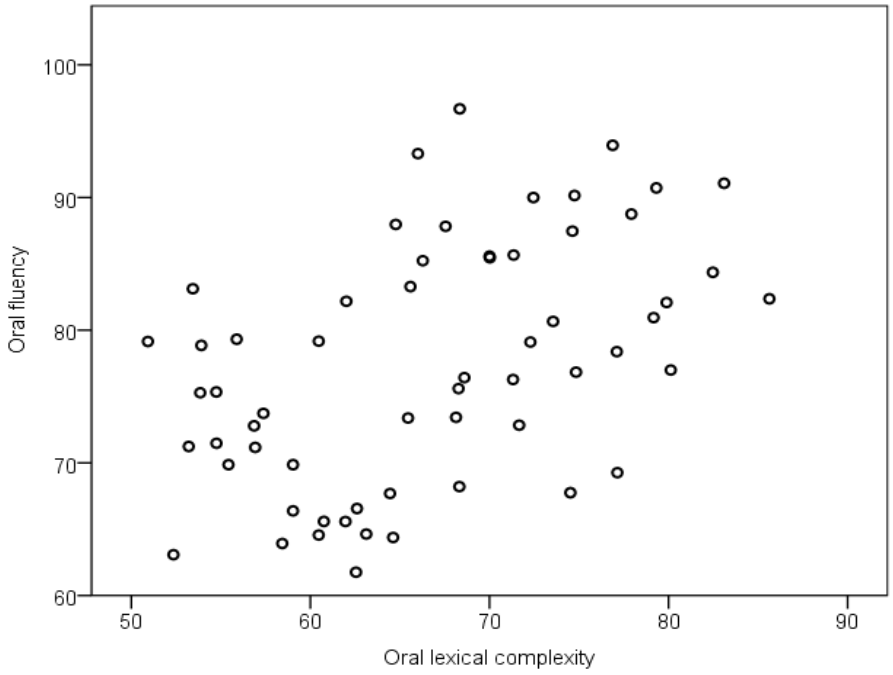

Figure 4. Scatter plot of the correlation between oral lexical complexity and oral fluency

As it is shown in the scatter plots, we can draw a straight line through the main cluster of the points showing a linear relationship. Therefore, the linearity assumption of the two pairs is not violated for running Pearson product-moment correlation coefficient. Furthermore, the shapes of the cluster are even from one end to the other in the scatter plot; so, our data enjoy the homoscedasticity assumption as well. Besides, in the scatter plot (Figure 4), the direction of the line that is drawn through the points, point from left to right upward in the scatter plot denotes a positive relationship: high scores on X, oral lexical complexity correlates with high scores on Y, oral fluency.

Table 8 represents the results of Pearson product-moment correlation among the three variables. As it is evident from Table 8, Pearson correlation found a significant positive relationship between oral grammatical complexity and oral fluency (r = .49, n = 60) at the significance level of .000 < .05, with high levels of oral grammatical complexity associated with high levels of oral fluency. Moreover, r squared was calculated as .24 via standard regression showing that 24% of the variance of oral fluency is

predicted by grammatical complexity. This effect size is small based on Cohen's guideline (1998, pp.79-81).

Table 8
Pearson Correlation for Oral Grammatical Complexity, Oral Lexical Complexity, and Oral Fluency Scores

		Oral fluency
Oral grammatical complexity	Pearson Correlation	.490**
	Sig. (2-tailed)	.000
	N	60
Oral lexical complexity	Pearson Correlation	.475**
	Sig. (2-tailed)	.000
	N	60

Also, as shown in Table 8, a significant positive relationship was detected between oral lexical complexity and oral fluency (r = .47, n = 60) at the significance level of .000 < .05, with high levels of oral lexical complexity associated with high levels of oral fluency. Besides, r squared was calculated as .23 using standard regression showing that 23% of the variance of oral fluency is predicted by lexical complexity. This effect size is small as well based on Cohen's guideline (1998, pp.79-81).

The Relationship between Fluency and Accuracy in Oral Performance

The third research question of this study aimed at investigating whether there is any significant relationship between fluency and accuracy in oral performance. Pearson product-moment correlation coefficient was performed to answer this research question. According to Pallant (2013, p. 133), "Pearson correlation is used to describe the strength and direction of the linear relationship between two continuous (interval) variables".

As it is illustrated in Table 9, One-Sample Kolmogorov-Smirnov Test revealed that scores of the two variables, that is, oral fluency (p = .90, p > .05) and oral accuracy (p = .85, p > .05) satisfy normality assumptions. Therefore, the parametric Pearson Product-Moment Correlation Coefficient was used to measure the relationship between the three main variables of the study.

Table 9
One-Sample Kolmogorov-Smirnov Test for Oral Fluency and Accuracy Scores

Variable	*N*	Kolmogorov-Smirnov Z	Sig.
Oral fluency	60	.500	.903
Oral accuracy	60	.569	.853

Before explaining the results of Pearson Product-Moment Correlation, the descriptive statistics for the scores of oral fluency and accuracy were computed and given in Table 10.

Table 10
Descriptive Statistics for Oral Fluency and Accuracy Scores

Variable	*N*	Mean	SD	Std. Error
Oral fluency	60	77.412	9.051	1.168
Oral accuracy	60	59.568	5.798	.748

Table 10 indicates the mean and standard deviation for the oral fluency scores ($\bar{x}$ = 77.41, *SD* = 9.05, *n* = 60) and for the oral accuracy scores ($\bar{x}$ = 59.57, *SD* = 5.80, *n* = 60). (See Appendix F for the raw scores gained on oral fluency and accuracy).

The scatter plot of the correlation between oral fluency and accuracy (Figure 6) was drawn before running correlation. As it is evident in the scatter plot, we can draw a straight line through the main cluster of the points signifying a linear relationship. Thus, the linearity assumption of the two pairs is met for performing Pearson product-moment correlation coefficient. Further, the shapes of the cluster are even from one end to the other in the scatter plot. Consequently, our data met the homoscedasticity assumption.

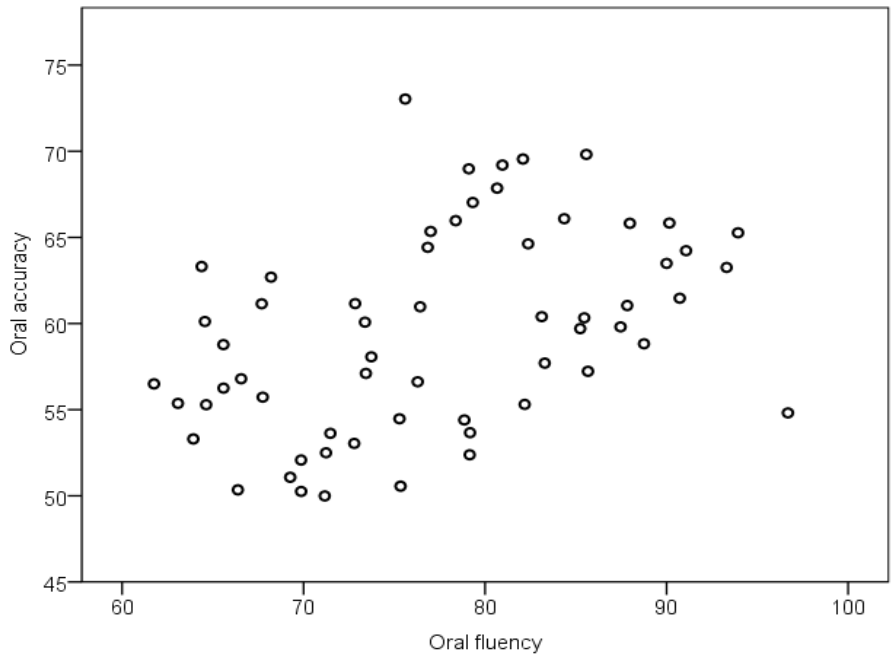

Figure 5. Scatter plot of the correlation between oral fluency and accuracy scores

In addition, the scatter plot (Figure 5) can tell us if the relationship between the two variables is positive or negative. In this scatter plot, the direction of the line that is drawn through the points, points from left to right upward in the scatter plot. This left to right upward trend reflects a positive relationship; high scores on X, oral fluency correlates with high scores on Y, oral accuracy.

Table 11 represents the results of Pearson product-moment correlation between the two variables.

Table 11
Pearson Correlation for Oral Fluency and Accuracy Scores

		Oral accuracy
Oral fluency	Pearson Correlation	.427**
	Sig. (2-tailed)	.001
	N	60

As it is evident from Table 11, Pearson correlation analysis revealed a significant positive relationship between oral fluency and accuracy (r = .43, n = 60) and a significant positive relationship between oral fluency and accuracy (r = .43, n = 60) at the significance level of .001 < .05, with high levels of oral fluency associated with high levels of oral accuracy. Further, r squared was calculated as .18 using standard regression showing that 18% of the variance of these two variables (i.e. oral fluency and accuracy) overlaps. This effect size is small as well based on Cohen's and Manion's guideline (Cohen & Manion, 1998).

Discussion

The research questions related to the relationship between proficiency and the CAF of speaking, and the data indicates that proficiency and spoken production are strongly correlated with each other in CAF. In addition, the relationships among CAF for speaking were highly correlated.

Considering the results of the statistical analysis for the relationship between three components of CAF in oral performance, there was a positive relationship between grammatical complexity and fluency, lexical complexity and fluency, grammatical complexity and accuracy, lexical complexity and accuracy and finally accuracy and fluency. In Skehan's Trade-off hypothesis when a component increases, it is expected that the other components decrease, something which the results of the current study do not support. To sum up, the results of this study ran against the Skehan's hypothesis.

In the same vein, the findings of the present study ran against the results of other studies like Ishikawa (2006). They argue that the results support Skehan and Fosters' (2001, p.193) preposition that "prioritization or predisposition (or both) seem to orient performance toward one (or two) of the three areas of accuracy, fluency, and complexity".

The findings of this study also differ from some previous studies on the relationship between accuracy and complexity. For example, Benevento and Storch (2011) observed learners' L2 writing assignments over a six-month period and reported significant improvements over time in language complexity and discourse, while accuracy did not show a distinct improvement. The variance seems related to the different in the participants'

proficiency levels; the proficiency level of the participants in the present study was intermediate (i.e., from beginning-intermediate to high-intermediate levels).

The findings of the study do not support what Freeman (2009) and Skehan (2003) believe in terms of limited attention resources either. They believe that one cannot attend to all aspects of language such as CAF features at the same time. Thus, there is a trade-off between these different aspects. In other words, if one pays attention to the fluency, he cannot dedicate the same level of attention to the accuracy or complexity as well. However, in the present study learners could pay attention to both accuracy and fluency.

The results of the study are consistent with those of Shiriyan and Nejadansari (2014) and Seyyedi, et. al. (2014) who showed that literature-based topics lead to more fluent and accurate second language oral production. However, such activities did not affect the text complexity. They argued that as these activities deal with emotion and feeling, learners have fewer problems in comparison to other activities. And as Skehan (2009) proposed tasks which are more familiar to the learners and whose structures are clear lead to higher accuracy and fluency than complexity. In this study, similar results were not obtained.

Previous research has emphasized that the three levels of proficiency may partially overlap and that complexity may decrease again at higher proficiency levels, as a result of task and genre effects, (Lu, 2011; Ortega, 2003; Pallotti, 2009; Polio & Yoon, this issue). Wolfe-Quintero et al. (1998) suggested that complexity measures might exhibit "omega-shaped" patterns, with an increase in complexity, followed by a decline at the higher proficiency levels. As argued by Ortega (2003), "more complex" does not necessarily mean "better," since a higher or lower complexity rate may also be determined by personal and stylistic choices, providing an indication of higher L2 proficiency. While the above-mentioned studies did not predict any correlation between the level of proficiency and oral CAF performance, in the present study advance learners' oral CAF was the highest and elementary learners' oral CAF was the lowest.

The findings of the study are in line with Ellis and Barkhuizen (Ellis & Barkhuizen, 2005) in that tasks with more cognitive demands push L2

learners to perform tasks in certain ways, prioritizing one or another aspect of language. Thus, complex tasks push learners to prioritize complexity over fluency. Tavakoli and Foster (2008) also argue that the more demanding a task in terms of its content, the more complex the language a learner will attempt performing a task. The explanation for the enhanced complexity may lie in the fact that complex tasks impose extra burden of information processing to the learners' mental capacities (Leaper & Riazi, 2014; Skehan, 2014, 2015).

To sum up, the obtained results showed that the findings of the current study reject Skehan's (2009) Trade-off Hypothesis, which means that when one component of CAF increased, the other components did not decrease and there was a correlation between all three components of CAF. Moreover, the initial the participants language proficiency level also played a crucial rule in their CAF performance, which means that learners with higher level of proficiency produced more complex, accurate, and fluent language.

To conclude, the present study was an attempt to find out the relationship between three components of CAF in oral performance at three proficiency levels. Considering the results of the data analyses, it was found that there was a significant positive relationship between oral grammatical complexity and oral accuracy and a significant positive relationship between oral lexical complexity and oral accuracy. Furthermore, a significant positive relationship between oral grammatical complexity and oral fluency and a significant positive relationship between oral lexical complexity and oral fluency were found out in this study. Additionally, the findings revealed a significant positive relationship between oral fluency and accuracy.

The findings of the current study can have implications for teacher assessors. Part of the role of instructors is to use assessment data so as to decide on teaching, instruction and students' learning. Teachers need to know how to analyze data based on students' knowledge, and decide what tasks learners should complete to indicate their competence. Teachers can manipulate the level of complexity of a task so that the task can be more suitable for learners' level of proficiency. For example, based on the findings of the current study, a task can be more complex if the teacher does not provide learners with the chance of thinking and planning. Meanwhile, it

is less likely for the teachers and course designers to have homogeneous classes and the teacher might find it difficult to have a task suitable for all learners, the findings of the present study indicate that teachers can add or remove one or two questions to or from a task to make it more complex or less complex for the learners with different level of proficiencies.

The work presented in this study on CAF presents new perspectives on the empirical study of CAF in SLA, and raises important theoretical and methodological questions including the need to further clarify testing instruments and better define the constructs to be measured as well as learner internal and external factors surrounding, affecting and impeding the development or manifestation of CAF in second language performance. These are all issues for further exploration. It is hoped that this study will contribute to further debate on CAF, shedding light on existing theoretical and methodological issues in the field as well as opening up new areas of inquiry.

This attempt mainly utilized a quantitative data collection, and did not focus on a qualitative account of the phenomena. The present study was exploratory by nature than explanatory and its goal was finding out the relationship rather than explaining such relationships. However, more expansions might be needed with the centrality of qualitative data analysis and discussion. Qualitative study will enrich the data and open grounds for future discussions and analysis.

In this study participants had similar L1 background. Will this study have the same results with participants of different L1 background? More similarities and differences may be discovered. However, this hypothesis will be left for future researchers to explore.

References

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35-59.

Bachman, L. F. (2004). *Statistical analyses for language assessment book*: Cambridge University Press.

Benevento, C., & Storch, N. (2011). Investigating writing development in secondary school learners of French. *Assessing Writing, 16*(2), 97-110.

Bradlow, A. R., Kim, M., & Blasingame, M. (2017). Language-independent talker-specificity in first language and second-language speech production by bilingual talkers: L1 speaking rate predicts L2 speaking rate. *The Journal of the Acoustical Society of America, 141*, 886–899.

Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses speed and repairs. *Language Testing, 30,* 159–175.

Cohen, L. M., & Manion, L. (1998). L.(1989) Research Methods in Education. *New York: Routledge*.

Ellis, R., & Barkhuizen, G. P. (2005). *Analysing learner language*: Oxford University Press Oxford.

Huensch, A., & Tracy–Ventura, N. (2016). Understanding second language fluency behaviour: The effects of individual differences in first language fluency, cross-linguistic differences, and proficiency over time. *Applied Psycholinguistics, 38*, 755– 785.

Huensch, A., & Tracy–Ventura, N. (2017). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *Modern Language Journal, 101*, 275–293.

Hunter, A.–M. (2017). *Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency*. (Unpublished doctoral dissertation). University of Surrey, Guildford, UK.

Ishikawa, T. (2006). The effect of task complexity and language proficiency on task-based language performance. *Journal of Asia TEFL, 3*(4).

Kim, Y., Nam, J., & Lee, S.-Y. (2016). Correlation of proficiency with complexity, accuracy, and fluency in spoken and written production: Evidence from L2 Korean. *Journal of the National Council of Less Commonly Taught Languages, 19*, 147-181.

Kormos, J., & Dénes, M. (2006). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*, 145–164.

Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.),

Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA (pp. 143–170). Philadelphia/Amsterdam: John Benjamins.

Leaper, D. A., & Riazi, M. (2014). The influence of prompt on group oral tests. *Language Testing, 31*(2), 177-204.

Mizera, G. J. (2006). *Working memory and L2 oral fluency*. University of Pittsburgh,

Pallant, J. (2013). *SPSS survival manual*: McGraw-Hill Education (UK).

Peltonen, P. (2018). Exploring connections between first and second language fluency: A mixed methods approach. *Modern Language Journal, 102*, 676– 692.

Robinson, P. (2003). The cognitive hypothesis, task design, and adult task-based language learning. *University of Hawai'I Second Langauge Studies Paper 21 (2).*

Robinson, P. (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (Vol. 2): John Benjamins Publishing.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*(4), 533-554.

Schmidt, R. (2001). Attention'in P. Robinson (ed.): Cognition and Second Language Instruction. *Cambridge, UK: Cambridge University Press*.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Segalowitz, N., French, L., & Guay, J.–D. (2017). What features best characterize adult second language utterance fluency and what do they reveal about fluency gains in short-term immersion? *Canadian Journal of Applied Linguistics, 20*, 90–116.

Seyyedi, K., Ismail, S., & Mohamed, A. (2014). The Effect of Task Structure on Second Language Learner's Narrative Writing Performance. *Journal of Foreign Languages, Cultures and Civilizations, 2*(1), 41-53.

Shiriyan, Z., & Nejadansari, D. (2014). The effect of literature-response activities on the complexity, accuracy, and fluency of Iranian EFL

learners' L2 oral productions. *Journal of Applied Linguistics and Language Research, 1*(2), 12-26.

Skehan, P. (1998a). *A cognitive approach to language learning*: Oxford University Press.

Skehan, P. (1998b). Task-based instruction. *Annual review of applied linguistics, 18*, 268-286.

Skehan, P. (2009a). Lexical performance by native and non-native speakers on language-learning tasks. In *Vocabulary studies in first and second language acquisition* (pp. 107-124): Springer.

Skehan, P. (2009b). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics, 30*(4), 510-532.

Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. *Processing perspectives on task performance, 211*, 211-260.

Skehan, P. (2015). Limited attention capacity and cognition. *Domains and Directions in the Development of TBLT, 8*, 123-155.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language teaching research, 1*(3), 185-211.

Tavakoli, P., & Foster, P. (2008). Task design and L2 performance. *Language Learning, 58*(2), 429-473.

Tavakoli, P., & Hunter, A.–M. (2018). Is fluency being "neglected" in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research*, *22*, 330–349.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied linguistics, 24*(1), 1-27.

Appendices

Appendix A

Percentage of Error-Free C-unit

No.	C-Unit	Error free C-unit
1	my father told to my teacher	0
2	when I was about ten years old	1
3	I used to cry so fast	1
4	a person say to me something	0
5	and umm I was offended	1
6	I'm very talkative	1
7	I wanted to be a doctor	1
8	now I like to be an English teacher	1
9	I see. I think um you have changed a lot	1
10	you became a patient person who is not talkative anymore and who is not crying for anything	1
11	do you think yourself that you have changed positive or negative	0
12	I think they are positive	1
	Total:	**9**
Error free C-unit/total number of C-unit * 100		
(9/12)*100 = 75%		

Appendix B

Clauses per C-unit

No.	C-Unit	Clause
1	my father told to my teacher	1
2	when I was about ten years old	1
3	I used to cry so fast	1
4	a person say to me something	1
5	and umm I was offended	1
6	I'm very talkative	1
7	I wanted to be a doctor	1
8	now I like to be an English teacher	1
9	I see. I think um you have changed a lot	3
10	you became a patient person who is not talkative anymore and who is not crying for anything	3
11	do you think yourself that you have changed positive or negative	2
12	I think they are positive	2
	Total:	**18**

Appendix C

Type Token Ratio

Rank	Word	Freq.	Rank	Word	Frq.	Rank	Word	Freq.
1	I	**10**	20	Very	**1**	39	Anymore	**1**
2	You	**5**	21	Became	**1**	40	Anything	**1**
3	You	**4**	22	That	**1**	41	Say	**1**
4	To	**4**	23	Or	**1**	42	About	**1**
5	A	**4**	24	Ten	**1**	43	Me	**1**
6	Was	**3**	25	Years	**1**	44	Do	**1**
7	Think	**3**	26	Old	**1**	45	Yourself	**1**
8	My	**2**	27	Told	**1**	46	Cry	**1**
9	When	**2**	28	Do	**1**	47	Negative	**1**
10	Talkative	**2**	29	Doctor	**1**	48	That	**1**
11	Have	**2**	30	Father	**1**	49	Crying	**1**
12	Who	**2**	31	English	**1**	50	Now	**1**
13	Teacher	**2**	32	Something	**1**	51	Like	**1**
14	But	**2**	33	Used	**1**	52	Wanted	**1**
15	Changed	**2**	34	Offended	**1**	53	‘m	**1**
16	Is	**2**	35	They	**1**	54	So	**1**
17	Be	**2**	36	Are	**1**	55	Patient	**1**
18	Not	**2**	37	An	**1**	56	Lot	**1**
19	Positive	**2**	38	Fast	**1**	57	See	**1** **Total 95**
Type-token ratio = (number of types/number of tokens)* 100								
=(57/95) *100 = 60%								

Appendix D

WPM

Number of token	Total task time in minute
96	**1.14**

Appendix E

Sample Tasks

Task 1

Task complexity: **low**

1. Many elements (more than two questions involved)
2. There-and-then (past tense use)

Procedure: Each group receives a handout (see below) of the answers and tries to work out the appropriate questions.

HOL I D A Y S
Ask your partner questions about his or her last holiday. Use the following notes to help you.
Where?
How long for?
Stay where?
With whom?
Like it? Why? Why not?
Sightseeing?
Sports?
Food?
Go again?
Do anything special?
Bad points?

Task 2

Task complexity: **mid**

1. Many elements (more than two questions)
2. There-and-then (past tense use)
3. Reasoning demand (provide reasons and why questions)
4. Without panning (no opportunity for the learners to plan their talk)

Students A and B: Talk to your each other about the best restaurant you have ever been to and ask the following questions.
Why was….?
Who …….?
What ….. eat? Why?
How much….?
Would you recommend……? Why?
Change the role

Task 3

Task complexity: **high**

1. Many elements (more than two questions involved)
2. There-and-then (past tense use)
3. Reasoning demand (provide reasons and why questions)
4. Without panning (no opportunity is provided for the learners to plan their talk)
5. Without prior knowledge (learners do not have prior knowledge about the topic)

Imagine you had a lot of brothers and sisters and you argued a lot. Talk about the most recent argument and say what you would argue about, who do you think would be guilty and why? How would you make up after the argument? And what should be done to prevent such arguments?

Biodata

Kazem Najmi is a PhD candidate in TEFL at Islamic Azad University, Science and Research Branch (IAUSRB), a university lecturer at IAUSRB and an English supervisor at a number of language centers. His main areas of research interest include task based language learning and learning through communication.

Masood Siyyari is an assistant professor of TEFL and applied linguistics at Islamic Azad University, Science and Research Branch (IAUSRB) in Tehran, Iran. He received his Ph.D. in applied linguistics from Allameh Tabataba'i University in Tehran, Iran. Currently, he supervises MA theses and PhD dissertations and teaches MA and PhD courses and applied linguistics and translation studies at SRBIAU. His main areas of research include language assessment, quantitative data analysis, and SLA.

Gholam-reza Abbasian is an associate professor of Applied Linguistics at Imam Ali University and Islamic Azad University, South Tehran Branch. He has presented papers at national and international conferences and is the author and translator of 15 books and many scholarly articles. He has been nominated a top scholar and teacher for seven consecutive years. He is also the internal manager of JOMM, one of the reviewers of Sage, FLA, and GJER journals, and a member of the editorial board of JSSIR.