Differential Time and Length Effect on EFL Learners' Oral Narrative CALF: Accounting for Variability

Omid Rezaei Dastgerdi[1*]

[1]Ph.D. in TESOL, Isfahan University, Isfahan, Iran

Citation
Rezaei Dastgerdi, O. (2025). Differential Time and Length Effect on EFL Learners' Oral Narrative CALF: Accounting for Variability. *International Journal of Language and Translation Research*, 5(2), pp. 1-19.

Available online

Keywords:

Accuracy,
Complexity,
Constraints,
Fluency,
Lexis,
Length,
Time

Abstract

The present study examined the differential effects of time and length constraints on complexity, accuracy, lexis, and fluency of EFL learners' narrative retellings. Thirty female participants at the intermediate proficiency level were selected non-randomly through a convenient sampling procedure from the researcher's intermediate level classes. They had to retell the narrative under time constraints and then reproduce it with length constraints. Following Tavakoli (2018), four indices (ratio of subordination, Weighted Clause Ratio, Diversity, and speech rate representing four dependent variables of complexity, accuracy, lexis, and fluency, respectively) out of 12 were chosen. Repeated-measures multivariate analysis of variance (MANOVA) in two performance sessions. Findings indicated that there was a difference in a linear combination of the four dependent variables of complexity, accuracy, lexis, and fluency across the two times of performance. It was also concluded that learners' narrative retellings do not demonstrate significant changes, in terms of complexity, accuracy, lexis, and fluency separately, across time-constrained versus length-constrained reproductions.

مطالعه حاضر اثرات متفاوت محدودیت‌های زمانی و طول گفتار را بر پیچیدگی، دقت، واژگان و روانی گفتار زبان‌آموزان زبان انگلیسی به عنوان زبان خارجی بررسی کرد. سی شرکت‌کننده زن در سطح مهارت متوسط به صورت غیرتصادفی و از طریق روش نمونه‌گیری در دسترس از کلاس‌های سطح متوسط محقق انتخاب شدند. آنها مجبور بودند روایت را تحت محدودیت‌های زمانی بازگو کنند و سپس آن را با محدودیت‌های طول بازتولید کنند. به پیروی از توکلی (۲۰۱۸)، چهار شاخص (نسبت تبعیت، نسبت وزنی بند، تنوع و سرعت گفتار به ترتیب نشان‌دهنده چهار متغیر وابسته پیچیدگی، دقت، واژگان و روانی گفتار) از بین ۱۲ شاخص انتخاب شدند. تحلیل واریانس چند متغیره با اندازه‌گیری مکرر (MANOVA) در دو جلسه اجرا. یافته‌ها نشان داد که در ترکیب خطی چهار متغیر وابسته پیچیدگی، دقت، واژگان و روانی گفتار در دو زمان اجرا تفاوت وجود دارد. همچنین نتیجه گرفته شد که بازگویی‌های روایی زبان‌آموزان، از نظر پیچیدگی، دقت، واژگان و روانی کلام، به طور جداگانه، در بازگویی‌های با محدودیت زمانی در مقابل بازگویی‌های با محدودیت طول، تغییرات معناداری را نشان نمی‌دهد.

[1] Corresponding Author's Email:
Omidrezaei.rezaei99@gmail.com

P-ISSN: 2750-0594
E-ISSN:2750-0608

Introduction

Investigating L2 discourse production appears to have implications for SLA (Yuan & Ellis, 2003). What has been intriguing researchers, and accordingly continues to stimulate literature, is the fact that the discourse as such demonstrates variability regarding fluency, complexity, accuracy, etc. Variability is accounted for within the cognitive approach to task, which was primarily a syllabus type, but over time, it turned into an area of research in its own right.

Task-prompted variability, on cognitive grounds, is attributed to humans limited short-term memory capacity. In fact, the attentional resources are supposed to be competitive so much so that aspects of tasks claim for attention differentially; thus, the L2 discourse production turns out to be subject to variations depending on task conditions, cognitive complexity, and difficulty. Aspects of variability are measured in terms of accuracy, complexity, and fluency. Planning is, indeed, one of the factors that can contribute to variation as it can probably determine task conditions, attentional cognitive resources, complexity, etc. In fact, there has been a growing interest in examining differential effects of planning time conditions on complexity, accuracy, and fluency (CAF), in English as a Second Language (ESL) and English as a foreign language (EFL) context. In what follows, the literature on planning is reviewed.

Review of Literature

Historically, Ellis (1987) was the first authority to come up with the idea that forms and functions which have not been completely automatized by the L2 learners are more accessible when planning takes place. He argued that opportunities for planned output are expected to increase the likelihood of these new, more difficult forms being eventually internalized and incorporated into the vernacular style, which engages in spontaneous production (Ortega, 1999, p. 111). Later, Ellis (2005) delineated that all the spoken and written language use, even the improvised and automatic language use, involves planning, and it is essentially a problem-solving activity because it presupposes deciding what linguistic devices need to be selected in order to affect the audience. Studies have shown that L2 learners produce more fluent and more complex language when they have the opportunity to plan a task before the performance (e.g., Foster & Skehan, 1996). Findings on planning (whether pre-planning or online) tend to contribute to variability (see Abdi & Ahangari, 2011; Logan, 2001, to name a few).

As Ellis (2005) suggests, planning has a ‘three-fold’ role to play. First, not only does it diminish online processing load, but it also reduces communication stress to achieve higher fluency. Second, planning helps students reach a high level of lexical and structural knowledge, which will enable them to produce more complex language. Third, it helps the conscious attention to form while it encourages the learners to produce more accurate language. According to Ellis (2005), there are two principal types of task-based planning: pre-task planning and within-task planning. Pre-task planning takes place before the task is performed, and within-task planning takes place during the performance of the task. Pre-task planning is further divided into rehearsal and strategic planning. Rehearsal planning is a practice performance that is done prior to the main

task performance. In other words, it involves repetition and practice as preparation for the real task. Strategic planning, on the other hand, involves learners preparing to perform the task based on the content they will need to encode and express. What distinguishes strategic planning from other types of pre-task planning activities is that the learners have access to the actual task materials. Likewise, Ellis divides within-task planning into two types: pressured and unpressured.

On the other hand, unpressured performance leads learners to engage in careful online planning (Ellis, 2005). While learners have a time constraint on their task performance in pressured task planning, there will be no such limit for task performance in unpressured task planning, which facilitates the way for planned language use (Ellis, 2009; Markee & Kunitz, 2013; Mytskowska-Wiertelak, 2011). Yuan and Ellis (2003), emphasize the advantages of on-line task planning over pre-task planning, which are (a) assisting students to search for the needed resources to encode the content, (b) helping students with controlling their output ahead of real-time production, and (c) backing the students up to review their output before producing it.

Following Ellis and Yuan (2006), pre-task planning promotes fluency and complexity but not necessarily accuracy in L2 learners' oral narratives. Although theoretical and empirical evidence highlights the importance of planning in oral performance, the role of pre-task and online planning in written production is still unclear (Ellis & Yuan, 2004). Clearly, different types of planning need to be systematically examined to show if they have any effect on the language produced in task performance. Lynch and McLean (2001) used a rehearsal with the learners of a medical profession who were learning an English for specific purpose course. They (ibid) reported that students with different proficiency levels appeared to benefit in different ways. Mostly, learners with a low proficiency level appeared to benefit in terms of accuracy and pronunciation, but those with a higher proficiency level took advantage of the opportunity to improve the clarity and economy of the explanations of a complex idea. Research on strategic planning has shown an effect on all three dimensions of production: complexity, accuracy, and fluency (Ellis, 2005). Several studies have pointed out that strategic planning enhances fluency. For instance, several studies have reported that giving students the opportunity of planning to students results in greater fluency (Ellis, 2005; Foster, 1996; Foster and Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997; Wendel, 1997; Yuan & Ellis, 1993).

Despite the positive role associated with planning, researchers are far from unanimous over what and how to measure speaking, which is probably one of the key challenges researchers face in the development of foreign language speaking. Specifically, in the Iranian context, the use of CALF is very rare among language teachers. It is widely admitted by several researchers that CALF helps us understand the relationship between linguistic output and second language underlying cognitive processes like attention and working memory (Robinson, 2011; Skehan & Shum, 2014). Moreover, it lets us operationalize underlying constructs of language production (Skehan, 2001). The research results confirmed promising results of using CALF to associate variation in syntactic complexity with variations in cognitive demands and even improved lexical diversity (Foster & Tavakoli, 2009; Kormos & Denes, 2004; Tavakoli & Foster, 2008). Although

studies that analyzed the effects of task planning and measuring production through CALF have been numerous, fundamental issues are left unanswered, such as the interaction between CALF components and different task conditions like time and length.

What has, however, gone largely unnoticed in literature is the fact that the very concept of 'planning' has hardly been explored in detail. Researchers tend simply to believe that allowing a time gap between a preparatory activity and the actual performance would necessarily guarantee engagement in planning on the individual's part. In other words, when the learner is allotted some time before the actual task performance, the planning has automatically taken place. Ellis (2006) maintains that.

> The teacher can elect to allow students to complete the task in their own time or can set a time limit. Lee (2000) strongly recommends that teachers set strict time limits. This option is important because it can influence the nature of the language that students produce. Yuan and Ellis (2002) found that giving students unlimited time to perform a narrative task resulted in language that was both more complex and more accurate in comparison to a control group that was asked to perform the same task under time pressure. The students used the time at their disposal to monitor and reformulate their utterances. Interestingly, the opportunity to plan online produced a different effect from the opportunity to engage in strategic planning, which led to greater fluency and complexity of language. It seems, then, that if teachers want to emphasize accuracy in a task performance, they need to ensure that the students can complete the task in their own time. However, if they want to encourage fluency, they need to set a time limit (Ellis, 2006, pp. 26-27).

Therefore, planning is viewed exclusively in terms of the time given. However, this can be inadequate in some ways. For one thing, planning may or may not take place if the time is allowed. Secondly, concerning within-task planning, one cannot make a clear-cut borderline between the learners' cognitive engagement in within-task planning on the one hand and actual performance on the other. Thirdly, it does not seem to be justifiable to reduce all planning down to a linguistic type. There is, conceivably, more to the story of planning than the linguistic aspects. The impact of implementation conditions (e.g., planning time conditions) and the impact of task design features (e.g., structured and unstructured) have been shown to prompts L2 learners to focus their attention on different aspects of language performance (Ellis & Yuan, 2004; Tavakoli & Skehan, 2005; Markee & Kunitz, 2013; Ahmadian, Tavakoli & Dastjerdi, 2015).

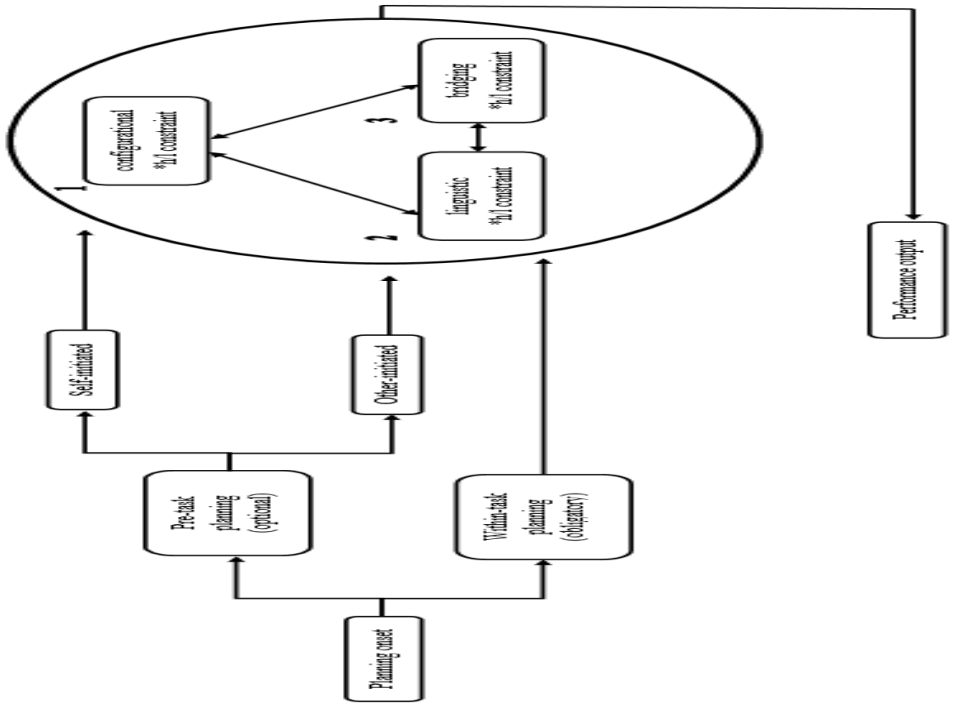
In the framework presented, planning start-up is launched as the initial step and is prompted by the test situation, teacher directions to start an activity or task, or even self-initiation, etc. Whatever the task happens to be, and however it is implemented in the context of learning, it triggers planning as a process, the beginning of which is the start-up stage. Having been launched upon by the learner, the planning, depending on task requirements, can occur before (i.e., pre-task planning) or during the actual task performance (within-task planning). In other words, pre-task

planning is not an essential feature of task performance, and it is cognitively possible to perform a task without planning online. Therefore, it is viewed as a non-obligatory aspect of the task. When launched upon, however, the pre-task planning is either self-initiated or other-initiated. By the former is meant the intention on the part of the learner to practice in advance of the actual performance without any demand from the teacher or others. Self-initiated pre-task planning is, for one thing, typified by the degree of constraint regarding time and length (configurational). An example can be a learner preparing for an oral presentation driven by his/her personal intention, which can be constrained to a certain extent (high or low) by time. Bridging is a mediating process whereby configurational demands (regarding time and length) are fashioned into linguistic units of production. It involves whether and how often and how long they have access to text, dictionary, task materials, etc.

Furthermore, the effect of rehearsal on different aspects of production has been examined. For example, Bygate (2001) investigated the effect of practicing specific types of tasks, involving narrative retelling and interviews, on both a second performance of the same task and performance of a new task of the same type. He reported that the second performance demonstrates greater complexity and fluency, and the opportunity given to learners to practice the particular task helps them extensively. However, there was no such practice effect in the performance of a new task of the same type.

Figure 1

Constraint-based plan framework for task-based performance



Accordingly, the following research questions are formulated.

RQ1. Do EFL learners' narrative retellings demonstrate significant changes in terms of a linear combination of complexity, accuracy, lexis, and fluency across time-constrained versus length-constrained productions?

RQ2. Do EFL learners' narrative retellings demonstrate significant changes in terms of complexity across time-constrained versus length-constrained productions?

RQ3. Do EFL learners' narrative retellings demonstrate significant changes in terms of accuracy across time-constrained versus length-constrained productions?

RQ4. Do EFL learners' narrative retellings demonstrate significant changes in terms of lexis across time-constrained versus length-constrained productions?

RQ5. Do EFL learners' narrative retellings demonstrate significant changes in terms of fluency across time-constrained versus length-constrained productions?

Method

Participants

The participants of this study were 30 sophomore female students majoring in English language and literature at the Islamic Azad University in Ardabil, Ardabil Province, Iran. They were doing their Oral Reproduction of Stories (ORS) course. All of the students shared the same L1 and the same level of proficiency. They were all placed at an intermediate level by a proficiency test at the beginning of the term. Not all of the participants completed the tasks during the data collection period. So the data reported here are from 26 participants performing the tasks at the two times of data collection on their final examination of the term. However, the sample size of the study, although not very large, is comparable to other studies in which CALF are used to examine interlanguage development, e.g., 28 in Feed, Segalowitz, and Dewey (2004), 39 in Leonard and Shea (2017), and 40 in Tavakoli (2018).

Materials and Instruments

Story No. 11 (see Appendix) from “Advanced Stories for Reproduction” (1965) was used in this study. It had a Flesch Reading Ease (FRE) index of 69.3. This passage was selected after employing the following steps. At first five books, which were mostly used in Iranian universities for the ORS course, were selected. Two stories from each book were randomly chosen. The mean FRE index for all ten passages was calculated. And at last, the above story was chosen because it had the nearest FRE index to the mean.

Procedure

The participants of the study were asked to meet at the language laboratory for their end-of-term exam in the ORS course. At first, copies of the passage were handed out among the students with ten minutes for them to read the text as many times as they could. Then all papers were collected. The participants were given four minutes for planning what to say. Subsequently, they were asked

to reproduce the story in 2:30 minutes. After this stage, they were given another four minutes for the planning and were asked to reproduce the text a second time, but just in five sentences. The first and the second reproductions were the applications of configurational constraints, i.e., high time and length demands, respectively.

All the allotted times (10 minutes for reading and two 4 minutes for planning) were calculated after a pilot study with five students of the same sex and language proficiency. The participants were not allowed to take notes or use any dictionaries during these steps because this was their final examination of the term. In the second step, where there was no time constraint, students were not allowed to leave the language laboratory right after they finished the task unless the last participant finished the recording of reproduction. The criteria for marking the second reproduction were the pauses at the end of each spoken sentence, which were taken as periods at the end of the written sentences. Both reproductions of all participants were recorded and transcribed, and the Complexity, Accuracy, Lexis, and Fluency (CALF) indexes were calculated.

CALF Measures

Based on Tavakoli (2018), in the current study, the following 12 indices were calculated. For complexity, length of AS (analysis of speech) unit, ratio of subordination, and length of clause were considered. Accuracy was measured by WCR (weighted clause ratio) and percentage of correct use of verbs, which stand for global and local accuracy, respectively. For lexis, D (diversity) and MTLD (measure of textual lexical diversity) were used to measure lexical diversity. The following indices were calculated for fluency: mean length of run, speech rate, number of silent pauses clause internal, number of silent pauses clause external, and composite repair measures (repetitions, hesitations, reformulations, and false starts).

Results

The study had a within-participants design with an independent variable with two levels of time and length constraints. The four dependent variables were CALF. The data consisted of the students' performances completed across the two main phases of the study. In each phase, the participants performed the retelling of the narrative individually.

Based on Tavakoli (2018), four measures, one from each category of CALF analysis (i.e., ratio of subordination, WCR, D, and speech rate), which were reported to represent CALF consistently, were selected from the total 12 indexes calculated in this study. Then, employing IBM SPSS (Version 28) software, a repeated-measures multivariate analysis of variance (MANOVA) was run to investigate whether there were statistically significant differences in learners' L2 proficiency in terms of CALF at the two times of data collection. At this level, all the prerequisites for running a MANOVA were observed.

At first, normal distribution was checked through the Kolmogorov-Smirnov test, and as it is presented in Table 1, three of the dependent variables (WCR, speech rate, and D) were normally distributed, and just one of the variables (subordination) did not reach the minimum significance.

Table 1

Test of Normality of Distribution

	Kolmogorov Smirnov		
	Statistics	df	Sig.
Mean subordination	0.201	26	0.008
Mean WCR	0.102	26	0.200
Mean D	0.157	26	0.099
Mean S rate	0.116	26	0.200

Before running any inferential analysis on data, descriptive statistics and an analysis of the design of the study were presented in Table 2. The numbers 1 and 2 refer to time and length of the study, respectively.

Table 2

Within-subject factors

Within-Subjects Factors		
Measure	Performance	Dependent Variable
ROS	1	Subordination1
	2	Subordination2
WCR	1	WCR1
	2	WCR2
D	1	D1
	2	D2
SR	1	Srate1
	2	Srate2

Note. Performance 1 indicates time-constrained and performance 2 indicates length-constrained task conditions. ROS = ratio of subordination; WCR = weighted clause ratio; D = diversity; SR speech rate.

As it is presented in Table 3, the standard deviation and mean statistics supported that disparity in performance for the variables of complexity and accuracy among the groups was somehow small and roughly identical for these two variables of the ratio of subordinates and weighted clause ratio. However, the mean differences regarding fluency and diversity of lexis were noticeable. The result of descriptive statistics revealed the mean scores of each pair across time and length.

Table 3 allows us to see that research question 1 regarding complexity is asking if 1.164 is significantly different from 1.117 across time and length, respectively. Research question 2 regarding accuracy is asking if 0.750 with the time variable is significantly different from 0.831

of the length variables. With the same token, research question 3 is asking if lexis knowledge differed significantly with the means of 63.594 and 76.740 across time and length. The last question asked if fluency represented a relatively different mean score over time and length (44.581, 53.336).

Table 3

Descriptive Statistics

	Mean	Std. Deviation	N
Ratio of Subordination (T)	1.164	0.200	26
Ratio of Subordination (L)	1.117	0.116	26
Weighted Clause Ratio (T)	0.750	0.134	26
Weighted Clause Ratio (L)	0.831	0.120	26
Diversity (T)	44.581	15.964	26
Diversity (L)	53.336	24.395	26
Speech Rate (T)	63.594	30.523	26
Speech Rate (L)	76.740	26.586	26

Note. N = number of participants. T = time-constrained task condition; L = length-constrained task condition.

In order to answer the research question number one, another multivariate test was run.

Table 4

Multivariate Tests

	Value	Hypothesis df	Error df	Sig.	Partial eta Square
Pillai's trace	0.379	3.354	22.000	0.027	0.379
Wilks' lambda	0.621	3.354	22.000	0.027	0.379
Hotelling's trace	0.610	3.354	22.000	0.027	0.379
Roy's largest root	0.610	3.354	22.000	0.027	0.379

RQ1: Is there any difference in a linear combination of the four dependent variables, ratio of subordination, WCR, speech rate, and D across the two times of performances?

As Table 4 shows, to answer the research question 1, multivariate tests can be applied. The results showed that a significant difference in the level of the linear combination of the four dependent variables, including ratio of subordination, WCR, speech rate, and D across the two times of performances equaled 0. 027. This is lower than .05, proposing that there is a linear combination of the four dependent variables across the two times of performance. Therefore, the null hypothesis of the study is rejected (Sig. = .027≤.05).

The importance of the impact of length and time on retellings of stories can be evaluated using the effect size statistic. Partial Eta Squared represents the proportion of the variance in the dependent variables (CALF) that can be explained by the independent variables of length and time constraints. The value, in this case, is 0.379. According to generally accepted criteria (Cohen, 1988), it is considered quite a large effect size. This implies 37.9% of the variance in the four dependent variables can be explained by the independent variable.

Since multivariate tests design involves a repeated measures component, the analysis needs the assumption of Sphericity. It holds the assumption that the variances of the differences between each level of the independent variable are not significantly different. Thus, the sphericity assumption needs to be met here for the independent variable of this study. Table 5 represented Mauchly's Test of Sphericity. However, this study investigated two levels of an independent variable, and the assumption of Sphericity cannot be tested as there are no degrees of freedom

Table 5

Mauchly's Test of Sphericity

Within	Measure	Mauchly's W	Approx. Chi-Square	df	Sig
Performances	ROS	1.00	0.000	0	0.00
	WCR	1.00	0.000	0	0.00
	D	1.00	0.000	0	0.00
	SR	1.00	0.000	0	0.00

RQ2: Do EFL learners' narrative retellings demonstrate significant changes in terms of complexity across time-constrained versus length-constrained productions?

According to the data provided in Table 6, there is no significant change in terms of complexity across time-constrained versus length-constrained productions. Since the significance value equaled 0.217 and this amount is larger than 0.0125 (after applying Bonferroni adjustment, 0.05/4 = 0.0125), this can be concluded that there is not any difference in terms of complexity across time and length constraints. Accordingly, the second null hypothesis of the study which proposed that EFL learners' narrative retellings do not demonstrate significant changes in terms of complexity across time-constrained versus length-constrained productions is confirmed (Sig. =.217>.0125).

RQ3: Do EFL learners' narrative retellings demonstrate significant changes in terms of accuracy across time-constrained versus length-constrained productions?

The significant amount of 0.013 was calculated which is bigger than 0.0125. This significance level suggests that there is not a difference between the subjects of the study regarding accuracy of their production over time-constrained versus length-constrained productions is accepted (Sig. = .013 > .0125).

RQ4: Do EFL learners' narrative retellings demonstrate significant changes in terms of lexis across time-constrained versus length-constrained productions?

The significance value for the lexis equaled .016. This amount is bigger than the critical value of .0125. Accordingly, the fourth null hypothesis that EFL learners' narrative retellings do not demonstrate significant changes in terms of lexis across time-constrained versus length-constrained productions is confirmed (Sig. = .016>.0125). As such, there was not a statistically significant difference in the lexis in regard to time factor compared with length factor in the production of the participants.

RQ5: Do EFL learners' narrative retellings demonstrate significant changes in terms of fluency across time-constrained versus length-constrained productions?

Table 6 revealed that the significance value for fluency equaled .072 which exceeded 0.0125. This result proposed that the fluency of production does not significantly change from time-constrained to length-constrained performances in this study. So the null hypothesis number 5 which proposed that EFL learners' narrative retellings do not demonstrate significant changes in terms of fluency across time-constrained versus length-constrained productions is confirmed (Sig. =.072 > .0125).

These results regarding the above-mentioned research questions and hypotheses are confirmed through running Pairwise Comparisons, which analyzed each dependent variable separately as no adjustment has been made for the fact that two contrasts have been made. According to Lynch et al. (2005), these results are the same as would be gained by analyzing each dependent variable separately like univariate test results.

Discussion

The finding of the study proposed that there was a difference in a linear combination of the four dependent variables, ratio of subordination, WCR, diversity, and speech rate of lexis across the two phases of performances. This is compatible with the findings of previous research which proposed the possible association between complexity, accuracy, lexis, and fluency (Mora & VallsFerrer, 2012; Witton-Davies, 2014). According to Sheakn and Foster (2012), it is proposed that there is an inter-relationship between CAF. As they justified, CAF sequencing is compatible with language development sequences which moves from complexity to try to gain control over form (accuracy) and finally leads to fluency. However, the performance regarding lexis was not included in their perspective.

Also, it is concluded that learners' narrative retellings do not demonstrate significant changes in terms of complexity across time-constrained versus length-constrained reproductions. This suggested that being engaged in retelling of stories does not impact complexity. The result was contrary to the study conducted by Tavakoli and Skehan (2005), who found that narrative retelling based on organized pictures led to higher complexity in participants' performance. It also contradicted the results in which the narrative tasks led to greater complexity, which suggested that learners used careful and extended language (Foster & Shekan, 1996). It is noteworthy that

most studies used picture and video-based narratives (Skehan & Foster, 1999). Some argued that the time pressure under which speaking is performed could be referred to as a hindrance in regard to syntactic complexity (Ellis, 2005; Kormos, 2006 & Skehan, 2009). On the other hand, several studies reported that complexity did not improve under short-term intervention (Gilabert 2007a; Kuiken & Vedder 2007a, b; Michel et al. 2007). This is totally compatible with the present study results.

The findings associated with accuracy were not improved under time and length restrictions of tasks. It is implied that the learners paid not enough attention to the accuracy during the retelling of stories under limited time and length task conditions. This result is contrary to several studies that have examined WCR to measure accuracy and found improvements in accuracy (Hong, 2007; Skehan & Foster, 2007; Tonkyn, 2007). It can be claimed that task conditions could have led to attention to form, but such attention is not as noticeable as predicted. The learners could have developed a higher accuracy awareness, which did not necessarily cause avoidance of errors (Skehan & Foster, 2012). They have continued that lower accuracy could have been devoted to syntactic choices. In general, as Tavakoli et al. (2016) argued, accuracy has shown limited development in short-term intervention. Bearing in mind that this study imposed a task time constraint, the insignificant accuracy in the present study can be attributed to time pressure. However, some argued that time limitation pressure could lead to greater accuracy (Yuan & Ellis, 2003).

Secondly, many have argued that less advanced learners are more involved in the monitoring process, which could reduce erroneous performance (Kormos, 2006; Levelt, 1989; Skehan, 2009a, 2014b). This involvement could work against accuracy. However, it needs to be mentioned that this study limited the scope of accuracy to the weighed clause ratio (WCR), while most research examined accuracy through rating WCR in three levels of clause boundary identification, categorization, and rating (Kuiken & Vedder, 2008). In spite of such different measurement level, it needs to be asserted that WCR, which is a clause-based syntactic unit, provides a more robust and in-depth distinction compared with the correct-incorrect classification in accuracy measurement (Bryfonski, 2020; Foster, Tonkyn & Wiggleworth, 2000).

Improvement in lexical diversity across task time versus length constraints did not occur. Research on systematic measuring of lexis has been rare (Skehan & Foster, 1996). Moreover, research concerning lexical measures has mostly examined two main areas of text-internal and text-external measures (Daller et al., 2003). It seems that lexical choices need more complex syntactic processing. The results concerning lexical diversity in this study do not akin to others who proposed that having planning time for both native and non-native speakers can result in producing different lexical choices (Shekan, 2009). Such insignificant lexical variation can be justified by the fact that seemingly second language speakers reveal lower abilities to integrate lexis, while this is not the case for native speakers (Skehan & Foster, 2012). On the other hand, attentional repertoires are limited (Cowan, 2005). While there is a trade-off between form and

meaning, meaning is mostly prioritized by the learners (Van Patten, 1990) and this could explain the less effort devoted to diversity in lexis.

Finally, the results were not in line with the findings from previous research. Length, speed of speech, and location of pauses, which accounted for fluency, did not differ significantly per participant across time and length task conditions. Contrary to my findings, many have argued that speech rate and mean length of run improve in different task conditions (Mora & Valls-Ferrer, 2012; Tavakoli et al., 2016). For instance, Tavakoli et al. (2016) stated that pedagogic intervention could help learners monitor their repetitions, false starts, reformulations, and hesitations during the performance, which leads to more fluency. Many others have reported that planning prior to task accomplishment led to greater fluency (Ellis, 2000; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997). Moreover, task conditions and planning time have been reported to improve the fluency of a second language over a short-term intervention (Bygate, 2001; Mehnert, 1998; Seifoori & Vahidi, 2012; Tavakoli & Skehan, 2005). It could be speculated from the results that fluency could have been sacrificed to produce more accurate utterances through pauses and monitoring (Kormos, 2006; Tavakoli, 2011). This could explain why the fluency did not differ significantly. However, my findings regarding fluency are compatible with other findings, which confirmed that learners paused more frequently, which lowered fluency, which could have been due to paying more attention to accuracy (de Jong et al., 2013; Kormos, 2006; Michel, 2011; Tavakoli et al., 2016).

Conclusion

In general, as the results of the present study offered, CALF may not be always a promising technique for improving communication. It is admitted that a successful verbal communication can take place without necessarily highly complex and accurate utterances which are diverse in lexis and are fluently delivered (de Jong et al., 2012; Kuiken & Vedder, 2012; Pallotti, 2009).

Regarding implications, in spite of the fact that each individual variable of CALF did not improve significantly over time and length task constraints in the present study, there are some implications. Theoretically, the findings of this study proposed that the application of CALF can be used both as a standard language performance measurement tool and as a suitable touchstone to examine language proficiency. This framework does not have the limitations that other measurements possess (Byfonski, 2020). This could be attributed to its simplicity in application and its sensitivity to minor changes in performance. It is also asserted that measures of task, including complexity, lexis, accuracy, and fluency, enable us to rate different features of performance comprehensively (Shekan, 2009). Pedagogically, although language learning in an EFL context has several limitations which are mainly related to input limitations, this shortcoming could be somewhat compensated by adjusting the instruction type, say, interventions. The result of the study confirmed that even within such a short-term intervention, significant gains in linear combinations of CALF were achievable, which is backed by several studies (Ellis, 2009; Tavakoli, 2018; Skehan, 2009).

The traditional classification of performance into correct and incorrect has several shortcomings. This makes sense, specifically regarding the oral data which could be the most difficult language skill to measure. Dealing with such data has its own issues. Researchers struggle as to how such data should be collected and analyzed and what units to be used to classify and measure them.

References

Abdi, M., & Ahangari, S. (2011). The effect of pre-task planning on the accuracy and complexity of Iranian EFL learners' oral performance. *Procedia – Social and Behavioral Sciences*, 29, 1950-1959.

Ahmadian, M., Tavakoli, M., & Dastjerdi, H.V. (2015). The combined effects of online planning and task structure on complexity, accuracy, and fluency of L2 speech. *The Language Learning Journal, 43*(1), 41- 56.

Boersma, P., & Weenink, D. (2007). *Praat version 4.5.01*. Computer software, downloaded from http://www.fon.hum.uva.nl/praat/

Bryfonski, L., & Ma, X. (2020). Effects of implicit versus explicit corrective feedback on mandarin tone acquisition in a scmc learning environment. *Studies in Second Language Acquisition, 42*(1), 61-88.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23-48). Longman.

Cowan, N. (2005). *Working memory capacity*, Psychology Press.

Crookes, G. (1989). Planning and interlanguage variation. Studies in Second Language Acquisition. *Cambridge University Press*, *11*, 367-383.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge University Press.

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition, 34*, 5-34.

De Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2013). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics. 10*.1017/S0142716413000210.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*, 12-20.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2005). *Planning and task performance in second language*. John Benjamins Publication.

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

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299-323.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics, 21*, 354-75.

Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and [+/- here-and-now]: Effects on L2 oral production. In M. P.

García Mayo (Ed.*), Investigating tasks in formal language learning* (pp. 44-68). Multilingual Matters. 98-116.

Freed, B., Segalowitz, N., & Dewey, D. (2004). Context of learning and second language fluency in French: Comparing Regular Classroom, Study Abroad, and Intensive Domestic Immersion Programs. *Studies in Second Language Acquisition, 26*(2), 275-301.

Givon, T. (1985). Function, structure and language acquisition. In *The cross-linguistic study of language acquisition*, 1, Dan Slobin (ed.), 1008-1025. Lawrence Erlbaum.

Graesser, A., McNamara, D., & Kulikowich, J. (2011). Coh-Metrix providing multilevel analyses of text characteristics. *Educational Researcher, 40*(5), 223-234.

Harper, D. (1986). *ESP for the university: ELT documents 123*. Pergamon Press.

Hong, R. (2007). Worry and rumination: differential associations with anxious and depressive symptoms and coping behavior. *Behavior research and therapy, 45*(2), 277-90.

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, *64*(4), 809-854.

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145-164.

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1), 48-60.

Leonard, K. R., &. Shea, C. E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity and underlying cognitive factors. *The Modern Language Journal, 101*(1), 179-193.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* The MIT Press.

Logan, K. (2001). The effect of syntactic complexity upon the speech fluency of adolescents and adults who stutter. *Journal of Fluency Disorders*. 26, 85-106.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. John Wiley & Sons.

Lynch, T., & Maclean, J. (2001). Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 141- 162). Longman.

Mackey, A. (1999). Input, interaction and second language development. *Studies in Second Language Acquisition, 21*, 557-587.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency

interviews using a new measure of lexical diversity. *Language Testing*, *19*(1), 85-104.

Markee, N. and Kunitz, S. (2013) Doing planning and task performance in Second Language Acquisition: An ethnomethodological respecification. *Language Learning*, *63* (4), 629-664.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-d, and HD-d: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*(2), 381-392.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83-108.

Michel, Marjie C., Kuniken, F., & Vedder, I. (2007). Effects of task complexity and task condition on Dutch L2. *International Review of Applied Linguistics, 45*, 241-59.

Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly, 46*, 610-641. doi:10.1002/tesq.34

Mystkowska-Wiertelak, A. (2016). Dynamics of classroom WTC: Results of a semester study. *Studies in Second Language Learning and Teaching*, *6*(4), 651-676. Newton, J. (2016). Researching tasks. *Language Teaching Research, 20*(3), 275-278.

Norris, J., Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments*. University of Hawai'i Press.

Norris, J., & Ortega, L. (2009). *Towards an organic approach to investigating CAF in instructed SLA*: The case of complexity. *Applied Linguistics*, *30*(4), 555-577.

Ochs, E., & Schieffelin, B. B. (1979). *Developmental pragmatics*. Academic.

Ong, J., & Zhang, L. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, *19*(2), 218-233.

Ortega, L. (1999). *Planning and focus on form in L2 oral performance. Studies in Second Language Acquisition.* Cambridge University Press, *21*, 10-14.

Perdue, C. (1993). Adult Language Acquisition. *Cross-linguistic perspectives, 2, the results*. Cambridge University Press.

Prabhu, N. (1987). *Second language pedagogy*. Oxford University Press.

Robinson, P. (2003). The Cognition Hypothesis, task design and adult task-based second language learning. *Second Language Studies*, *21*(2), 45-105.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43*, 1-32.

Robinson, P. (2011). *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance.* John Benjamins.

Sato, C. (1990). *The syntax of conversation in interlanguage development*. Gunter Narr.

Schank, R., Bermanm, T., & Kimberli Macpherson, K. (1999). Learning by doing. In

Instructional-design theories and models: a new paradigm of instructional theory, Volume II, Charles Reigeluth (ed.), 161-182. Erlbaum.

Schmidt, R. W. (1995). Consciousness and Foreign Language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-65). University of Hawaii Press.

Schmidt, R. W. (2001). Attention. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 3-30. Cambridge University Press.

Seifoori, Z., & Vahidi, Z. (2012). The impact of fluency strategy training on Iranian EFL learners' speech under online planning conditions. *Language Awareness, 21*(1), 101-112.

Skehan, P. (1996). Second Language Acquisition Research and Task-based Instruction. In J. Willis and D. Willis (Eds), *Challenge and Change in Language Teaching* (pp. 17-30). Oxford Heinemann.

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

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*(3), 1-27.

Skehan, P. & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In Van Daele, S., Housen, A., Kuiken, F., Pierrard, M. & Vedder, I. (Eds.), *Complexity, Accuracy, and Fluence in Second Language Use, Learning, & Teaching.* pp. 207-226. Contactforum.

Skehan, P. & P. Foster. (2012). Complexity, accuracy, fluency and lexis in task-based performance. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 199-220. John Benjamins.

Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 197- 210). John Benjamins.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th.eds). Pearson Education.

Tavakoli, P. (2018). L2 development in an intensive Study Abroad EAP context. *Elsevier Ltd, 72*, 62-74.

Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Q, 50*, 447-471.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439- 473.

Tavakoli, P. & Skehan, P. (2005). Planning, task structure, and performance testing In R. Ellis (Ed.), *Planning and Task Performance in a Second Language.* John Benjamins.

Tonkyn, A. (2007). Short-term changes in complexity, accuracy and fluency: Developing progresssensitive proficiency measures. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use,*

learning and teaching (pp. 263-284). KVAB.

Van Patten, B. (1990). Attending to content and form in the input: an experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

Wendel, J. (1997). *Planning and second language narrative production. Unpublished doctoral dissertation*, Temple University Japan, Tokyo.

Witton-Davies, G. (2014). The study of fluency and its development in monologue and dialogue. *Unpublished doctoral dissertation.* University of Lancaster.

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

Appendices

Appendix A

Passage for Reproduction

Mr and Mrs Davies had left their Christmas shopping very late. There were only a few days more before Christmas, and of course the shops and streets were terribly crowded, but they had to get presents for their family and friends, so they started out early one morning for the big city, and spent several tiring hours buying the things they wanted in the big shops.

By lunch-time, Mr Davies was loaded down with parcels of all shapes and sizes. He could hardly see where he was going as he and his wife left the last shop on their way to the railway station and home. Outside the shop they had to cross a busy street, made even busier than usual by the thousands of people who had come by car to do their last-minute Christmas shopping.

Mr and Mrs Davies had to wait for the traffic lights to change, but as Mr Davies could not see in front of him properly, he gradually moved forward into the road without realizing it. Mrs Davies saw this and became worried. Several times she urged her husband to come off the road, but without success. He could not hear her because of the noise of the traffic.

Finally, she shouted in a voice that could be heard clearly above all the noise, 'Henry! If you intend to stand in that dangerous position a moment longer, give *me* the parcels!'

Appendix B

Sample reproductions of two students under time and length constraints respectively

B.1 Time-Constrained Performances

Student 1:

In the name of God this story is about the M… Mr and Mrs James that they lived the very early the sh… Christmas shopping… They went the …s… city for buying buying the m… Christmas tools but they they left out their their shopping very early morning and … the … streets was very crowded when they reached the station railroad station and reached the and they moved to city or wherever reached to city. Mr Chames Mr Jims moved very for … and very fast a… and than his wife and Mr Mrs Miss raised voiced. his wife bad but he didn't answered he her because he couldn't hear his voice because of the higher voice of the traffic, but when….

Students 2:

Mr Davi— Mr and Mrs Davis had left their Christmas shopping very la— very late. She they went to the to the shopping because they because they because they had to get present for their family and friends and sta— they started at started out started at early morning to big city because they want went to shopping from the big sg… big shop big shop. By lunchtime, Mr Davis was loaded down with the parcels and he can't parcel of all of the shapes and sizes. She then he cannot see anywhere anywhere anywhere of him anywhere he was going and the street was busier than usual by the thousand people… H… he and they had to wait for the traffic light and but Mr Davis can't see in front of him and he gradually leaved of forward without realized it and he Mrs Davis———.

B.2 Length-Constrained Performances

Student 1:

Mr James and his wife have Christmas shopping. They one day they left had left her shopping were left to the city you know you know you know you know you know I mean the city because of the city was very busy a… because of the Christmas and … in the city his wife her wife his wife wants to heard him but traffic voice doesn't let him heard her voice and the finally he heard his vo… her v… her wife voice because because of her voice because because higher voice of the than usual in traffic voice.

Student 2:

MR and Mrs Davis had left their home for Christmas shopping and they had they had they had to get present for fa… and for their family and friends. By the lunchtime when they out of the shop. Mr Davis was load down with the parcels and he cannot he could not see anywhere he was going and the street was very busy and they had to wait for traffic light to change and he couldn't see in fron of him and he gradually moved he gradually moved you know forward without realized it. Mrs Davis tried to say told him said to him but he can't he and he couldn't hear the voice of him the voice of her. Finally, she finally he could hear the r… the her ro… finally she he could hear her voice very clearly. He she said she said she said if you if you if you don't stand danger in danger position, I give I give you the they parcel.