

Lexical Complexity as a Function of Task Type and Proficiency Level in the Speech Monologs of Iranian EFL Learners

Aso Bayazidi¹, Ali-Akbar Ansarin², Zhila Mohammadnia^{3*}

¹ Ph.D. Candidate, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran, *bayazidi.aso@gmail.com*

² Professor, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran, *ansalak@gmail.com*

^{3*} Assistant Professor, Department of English Language and Literature, Faculty of Literature and Humanities, Urmia University, Urmia, Iran,
z.mohammadnia@urmia.ac.ir

Abstract

Despite the abundance of research evaluating the effects of task complexity, task types, and planning on the complexity, accuracy, and fluency of the language produced by learners, most studies have focused on the syntactic aspect of complexity, with very few studies investigating the lexical part of complexity. Such studies explored the lexical performance of learners through using merely one measure of lexical complexity (namely diversity). The present study is an attempt to further explore the effects of task type and proficiency level on different aspects of lexical complexity of spontaneous speech monologs produced by intermediate and advanced Iranian EFL learners. To this end, 35 intermediate and advanced Iranian learners of English performed three different speaking tasks: an argumentation, a description and a narration task. The speech monologs were coded for three different aspects of lexical complexity: diversity, density and sophistication. Two-way repeated-measures ANOVAs were conducted to evaluate the main effects of task type (the within-subjects variable) and proficiency level (the between-subjects variable) on lexical complexity. The results showed that task type and proficiency level both significantly affect lexical complexity in the participants' task performance. The argumentation task yielded the highest scores for diversity and density, while the highest sophistication score was obtained for the narration task. There was no interaction between task type and proficiency level for the diversity and sophistication scores, and the advanced learners consistently got higher diversity and sophistication scores for all the three tasks, while there was an interaction between the two variables for the density scores; the advanced learners got higher density scores for the description and narration tasks but not on the argumentation task.

Keywords: Density, Diversity, Lexical Complexity, Sophistication, Task Type

Received 25 August 2019

Accepted 14 November 2019

Available online 18 December 2019

DOI: 10.30479/jmrels.2019.11406.1418

©2020 by the authors. Published by Imam Khomeini International University.



1. Introduction

Researchers in the areas of applied linguistics and instructed second language acquisition have always been interested in measuring second language performance. Brumfit (1979) proposed accuracy on the one hand and fluency on the other hand as two important aspects of language use. Skehan (1998) added complexity as another important aspect of language use, and thereby the triad of complexity, accuracy, and fluency (CAF) were introduced as the three fundamental dimensions that characterize second language performance. These three components have proved useful measures of second language performance (Skehan, 2009b).

In the past two decades, investigating the effects of such independent variables as task complexity, task type, task repetition, and planning on the complexity, accuracy, and fluency of second language learners' linguistic performance on pedagogical tasks has been a thriving area of research. Ellis (2003) offers the following working definitions for the three dimensions. Complexity refers to the extent to which the language produced by the learners is elaborate and varied. It is divided into syntactic and lexical complexity. Accuracy is defined as the extent to which the language produced by the learner conforms with target language norms. Fluency refers to the extent to which the language produced by the learner manifests pausing, hesitation, or reformulation. Skehan (2009b) characterizes successful task-based performance as containing "more advanced language, leading to complexity; a concern to avoid errors, leading to higher accuracy if this is achieved; and the capacity to produce speech at normal rate and without interruption, resulting in greater fluency" (p. 510).

In terms of the cognitive underpinnings of CAF, complexity and accuracy are associated with the current state of the learner's L2 knowledge representation and restructuring, while fluency is related to control and automatization of L2 knowledge (Housen et al., 2012; Skehan, 2009b). Two competing models have been proposed to account for the complexity, accuracy, and fluency of L2 learners' production in task performance. The Trade-off Hypothesis (Skehan, 1998, 2003) argues that humans have a limited processing capacity and attending to one dimension of language production may take away attention from others. According to the Trade-off Hypothesis, raised performance in one dimension may be achieved at the expense of performance in other dimensions. This competition shows itself most prominently in the prioritization of meaning (complexity) over form (accuracy) in tasks that are cognitively more demanding. In contrast, the Cognition Hypothesis (Robinson, 2001, 2003, 2005) argues that human attention resources are multiple and that speakers have the capacity to handle different demands on their attention simultaneously. Therefore, complexity and accuracy can go together. Testing these two rival models has proven

difficult, in part due to the lack of conceptual and operational clarity of the dependent variables (Housen et al., 2012). As a result, the results of empirical studies so far have not been consistent and do not equivocally support either of the two models (Robinson, 2011; Robinson & Gilabert, 2007; Skehan, 2009b).

One shortcoming of the research done so far on task performance is that it has focused mostly on the syntactic aspect of complexity, with very few studies investigating the lexical aspect of this performance area. Skehan (2009b) states that lexis has been strikingly absent in task research and that it is vital to incorporate some measures of lexis into task performance. The three dimensions of complexity, accuracy, and fluency thus need to be supplemented by measures of lexical performance. The range of measures also needs to be widened to cover this additional area. Also, most of the studies conducted in the area of task performance have used only lexical diversity as the measure of lexical performance.

2. Literature Review

The complexity component of the CAF triad is divided into lexical and syntactic complexity. Lexical complexity is a multidimensional feature of language use which encompasses diversity, sophistication, and density (Read, 2000; Wolfe-Quintero et al., 1998). Research into lexical measures also makes a distinction between text-internal and text-external measures (Daller et al., 2003). The text itself is enough to calculate text-internal measures, while text-external measures require some sort of reference material, which are usually based on word frequency lists. Lexical diversity is an example of text-internal measures, which has traditionally been measured through some sort of type-token ratio (TTR). A serious problem with TTR measures is that they are affected by text-length or sample size and a correction has to be made (Malvern & Richards, 2002). A general accepted measure of lexical diversity developed to address the sensitivity to text length is D (Malvern & Richards, 2002; Malvern, Richards, Chipere, & Durán, 2004; Richards & Malvern, 2007), which is calculated by the VOCD subroutine within Computerized Language Analysis (CLAN) (MacWhinney, 2000). For the present, D seems to be the best measure of lexical diversity (Jarvis, 2002; McCarthy & Jarvis, 2007). Research has suggested that lexical diversity distinguishes between writing tasks with differing features such as genre (Olinghouse & Wilson, 2013; Yoon & Polio, 2017) and task complexity (Frear & Bitchener, 2015; Kuiken & Velder, 2008). Researchers have also suggested that lexical diversity contributes to lexical proficiency scores (e.g., Bulté & Housen, 2012; Crossley, Salsbury, McNamara, & Jarvis, 2011). One may ask what such a measure measures. “D provides an index of the extent to which the speaker avoids the recycling of the same set of words.

If a text has a lower D, it suggests that the person producing the (spoken or written) text is more reliant on a set of words to which he or she returns often.” (Skehan, 2009a, p. 108).

In contrast, lexical sophistication is argued to involve the depth and breadth of vocabulary knowledge (Read, 2000) and has been measured through corpus-based frequency lists and range scores. Corpus-based measures take frequency lists from corpus analysis and then compute how many words defined as difficult are used in a text, with difficulty being defined on the basis of lower frequencies. Laufer and Nation’s (1999) Lexical Frequency Profile is the most well-known measure of this sort. The profile provides information on the number of words in a text drawn from the 1000 word level, the number drawn from the 2000 word level, and so on. It enables a judgment to be made regarding the extent to which very frequent words are relied upon less. An alternative measure is P_Lex developed by Meara and Bell (2001), which uses a mathematical modeling procedure. It divides a text into ten-word chunks and computes the number of infrequent words in each ten-word chunk, as measure named lambda. Additionally, average word length has been used by some researchers (e.g., Jarvis, Grant, Bikowski, & Ferris, 2003) as a measure of lexical sophistication based on the assumption that longer words tend to be more sophisticated. However, Verspoor, Schmid, and Xu (2012) questioned the validity of average word length as a measure of lexical sophistication since it neither predicted overall writing quality well nor effectively discriminated between proficiency levels.

Lexical density is defined as the ratio of lexical words (or content words) to the total number of words (Ure, 1971). Lexical words include nouns, adjectives, verbs, and adverbs and give a text its meaning and provide information regarding what the text is about. Other kinds of words such as articles, prepositions, and conjunctions are more grammatical in nature and give little or no information about what a text is about. These non-lexical words are called function words. Lexical density is simply a measure of how informative a text is. Spoken texts tend to have a lower lexical density than written ones (Halliday, 1985).

Using the same research design, Foster and Skehan (1996) and Foster (2001) investigated the effects of three different types of speaking tasks (a personal, a narrative and a decision-making task) on two aspects of lexical complexity (lexical diversity and lexical sophistication). The results indicated that narrative tasks consistently produce the highest sophistication values (as measures by lambda). Regarding lexical diversity, the personal tasks were inconsistent, but the pattern of results for the narrative and decision-making tasks are the reverse of those for lexical sophistication. The decision-making tasks consistently yielded higher diversity figures (as measured by D) than the narrative tasks.

Yoon and Polio (2017) examined the linguistic development of ESL students across two written genres. They used average word length and the word frequency as measures of lexical sophistication, and lexical diversity was measured through vocd-D. The results indicated a significant genre effect with very large effect sizes for both measures. The two measures, however, behaved in different ways. While the argumentative essays resulted in greater lexical sophistication, narratives showed greater lexical diversity.

Using the EF-Cambridge Open Language Database (EFCAMDAT), Alexopoulou, Michel, Murakami, and Meurers (2017) explored the effect of selected task types (narrative, description, and professional tasks) on the complexity and accuracy of the language they elicit. They used the Measure of Textual Lexical Diversity (MTLD; McCarthy & Jarvis, 2010) to measure global lexical complexity. The results showed that professional tasks (the tasks requiring the participants to write a resume and a job advertisement) resulted in the highest diversity scores, as compared with the description and narrative tasks.

As was mentioned earlier, despite the abundance of research done on the effects of such independent variables as task complexity, task types, and planning on the complexity, accuracy, and fluency of the language produced by learners, most studies have focused on the syntactic aspect of complexity, with very few studies investigating the lexical part of complexity. Also, those studies that have explored the lexical performance of learners have used only one measure of lexical complexity (mostly diversity). The present study is an attempt to further explore the effect of task type as an important factor on the lexical complexity of spontaneous speech monologs produced by intermediate and advanced Iranian EFL learners. The research questions that guide the present study are as follows:

1. How does L2 learners' lexical performance differ across different task types?
2. How does L2 learners' lexical performance differ across different proficiency levels?

3. Method

This study had a 2×3 mixed repeated measures design with two independent variables. These independent variables were the within-subjects variable of task type having three conditions (argumentation, description and narration) and the independent between-subjects variable of proficiency level having two levels (intermediate and advanced). The dependent variables were lexical diversity, lexical density and lexical sophistication.

3.1. Participants

The participants in the study were 18 intermediate (55.55% male) and 17 advanced (70.58% male) learners of English as a foreign language (EFL) in the adults' department at the Iran Language Institute (ILI) from two intact classes who participated in the study voluntarily. They ranged in age from 18 to 32 (mean = 21.97, SD = 3.75). The intermediate participants had been studying English for 10 semesters and the participants in the advanced group had been studying for 18 semesters in the institute's regular classes, which run two sessions a week, each session lasting two hours. Each term lasted for 10 weeks, and the participants had to take the ILI's final exam which tested all the four main language skills at the end of each term. Also, the participants' speaking ability was regularly evaluated during the semester by the teacher. In addition to studying English at the private language institute, the participants had studied English as a school subject two hours a week for six years in the Iranian national education system. The participants came from a variety of L1 backgrounds (Azerbaijani, Turkish, Kurdish, and Persian). None of the participants had ever lived or stayed in an English-speaking country. All the participants in the study signed an informed written consent form.

3.2. Tasks

Three tasks were used to elicit spontaneous speech monologues from the participants. In the argumentation task, the participants were asked to respond to the question whether money can make people happy. The description task required the participants to describe someone they enjoyed spending time with. In the narration task, the participants were first asked to inspect a series of cartoon pictures with no text and then to narrate a story based on the pictures.

3.3. Measures

Lexical diversity: D was used as a measure of lexical diversity, which was calculated using the VOCD subprogram within Computerized Language Analysis (CLAN) (MacWhinney, 2000) available at www.textinspector.com.

Lexical density: Lexical density is defined as the number of lexical words divided by the total number of words. The web-based Lexical Complexity Analyzer developed by Xiaofei Lu (Ai & Lu, 2010; Lu, 2012) available at www.aihaiyang.com/software/lca was used to calculate lexical density.

Lexical sophistication: Lexical sophistication is operationalized as the percentage of words beyond the 2000 most frequent English words based on Corpus of Contemporary American English (COCA) frequency lists. The lexical tools available at www.textinspector.com were used to calculate the

number of types beyond the 2000 most frequent words. The number was then divided by total number of types to obtain the lexical sophistication measure.

3.4. Procedure

Each participant was interviewed individually. The participants' responses were recorded using a digital voice recorder. The three tasks were presented to the participants in a counterbalanced order. For each task, the participant was given 30 seconds to plan his/her response, during which time they were not allowed to take notes. The recorded performances were then transcribed as Word documents. The transcriptions were pruned by removing false starts, repetitions, and self-corrections. The pruned transcriptions were then coded for the three measures that were used to operationalize the different aspects of lexical complexity in the study.

4. Results and Discussion

4.1. Results

This study was aimed to investigate the possible effects of task type and proficiency level on various dimensions of lexical complexity in the monologs of Iranian EFL learners. Table 1 shows the descriptive statistics for the mean diversity, density and sophistication scores across the three different tasks. As can be seen, the highest diversity score was obtained for the argumentation task, followed by the description and narration tasks, respectively. The same pattern of results was found for the density aspect of lexical performance; the argumentation task yielded the highest density score, followed by the description and narration tasks. As for the sophistication aspect of lexical performance, it was the narration task that showed the highest score, followed by the argumentation and description tasks, respectively. Therefore, the argumentation task resulted in the highest diversity and density scores, while the highest sophistication score was obtained for the narration task.

4.1.1. The Effect of Task Type on Lexical Complexity

The first research question addressed the effect of task type on different aspects of lexical complexity. To see if the difference in the scores obtained for the three tasks were statistically significant, a mixed repeated measures analysis of variance (ANOVA) was conducted for each aspect of lexical complexity. The results of the three ANOVA tests for the three different aspects of lexical complexity are shown in Table 2. The sphericity assumption was not violated in any of the three tests. As can be seen, task type affects all the three different aspects of lexical complexity in a significant way. Therefore, there is a main effect of task type in the

participants' lexical performance on all the three aspects of lexical complexity.

Table 1

Descriptive Statistics for the Mean Diversity, Density and Sophistication Scores in Three Tasks

		Argumentation		Description		Narration	
		Mean	SD	Mean	SD	Mean	SD
Diversity	Adv. (N=17)	59.42	8.99	58.97	13.04	40.33	10.78
	Int. (N=18)	49.72	12.51	49.36	10.16	33.18	6.05
	Total (N=35)	54.43	11.85	54.03	12.47	36.66	9.28
Density	Adv. (N=17)	0.45	0.032	0.43	0.033	0.45	0.030
	Int. (N=18)	0.46	0.033	0.41	0.030	0.42	0.025
	Total (N=35)	0.46	0.032	0.42	0.032	0.43	0.033
Sophistication	Adv. (N=17)	0.144	0.080	0.166	0.087	0.228	0.102
	Int. (N=18)	0.088	0.046	0.093	0.043	0.161	0.068
	Total (N=35)	0.116	0.070	0.128	0.076	0.194	0.091

Table 2

Two-Way Repeated-Measures of ANOVA for the Main Effect of Task Type on Three Aspects of Lexical Complexity

	Mauchly's Test		Repeated Measures ANOVA			
	Mauchly's W	Sig.	F	Sig.	Effect Size (η^2)	Observed Power
Diversity	.872	.111	43.802	.000	.570	1.00
Density	.988	.823	18.600	.000	.360	1.00
Sophistication	.989	.839	19.093	.000	.367	1.00

Table 3 shows the results of pairwise comparisons (Tukey post hoc tests) of the three different tasks for each of the three different aspects of lexical complexity. As for the diversity aspect, the difference in mean scores between the argumentation and description tasks was not significant, while the two tasks both yielded scores that are significantly higher than the

narration task. Each pairwise comparison was significant for the density aspect of lexical complexity. The pattern of results found for the sophistication aspect of lexical complexity was opposite the one obtained for the diversity aspect. The argumentation and description tasks both yielded sophistication scores that are significantly lower than the narration task, while the difference between the two tasks themselves is not significant.

Table 3

Pairwise Comparisons of the Tasks for Three Aspects of Lexical Complexity

		Mean difference	Std. error	Sig.
Diversity	Argumentation vs. Description	0.407	2.52	0.873
	Argumentation vs. Narration	17.81	2.06	0.000
	Description vs. Narration	17.40	1.87	0.000
Density	Argumentation vs. Description	0.041	0.006	0.000
	Argumentation vs. Narration	0.024	0.007	0.002
	Description vs. Narration	-0.017	0.007	0.013
Sophistication	Argumentation vs. Description	-0.013	0.013	0.326
	Argumentation vs. Narration	-0.078	0.014	0.000
	Description vs. Narration	-0.065	0.013	0.000

We were also interested in any possible interactions between task type and proficiency level in the participants' lexical performance on the three tasks. The interaction between task type and proficiency level was not significant for diversity ($F(2, 66) = .222, p = .801, \text{effect size} = .007, \text{observed power} = .083$) or sophistication ($F(2, 66) = .209, p = .812, \text{effect size} = .006, \text{observed power} = .081$) scores. As can be seen in Figures 1 and 2, the advanced learners consistently did better than the intermediate learners in terms of how diverse and sophisticated the words they used were across the three tasks.

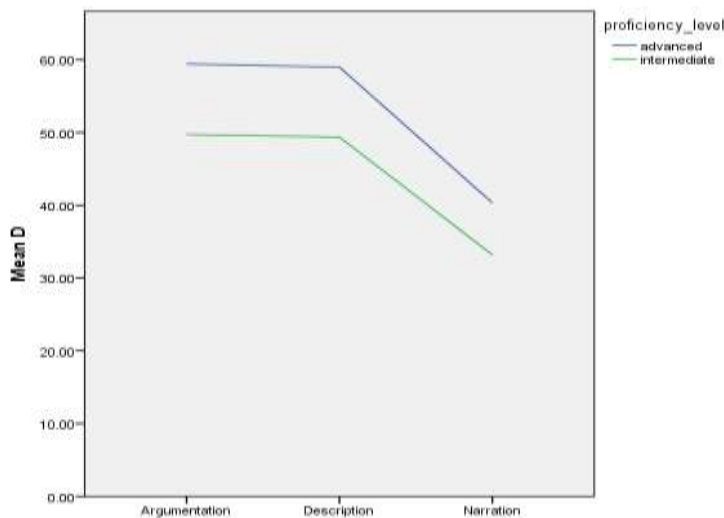


Figure 1. Intermediate and Advanced Learners' Mean Diversity Scores across the Three Tasks

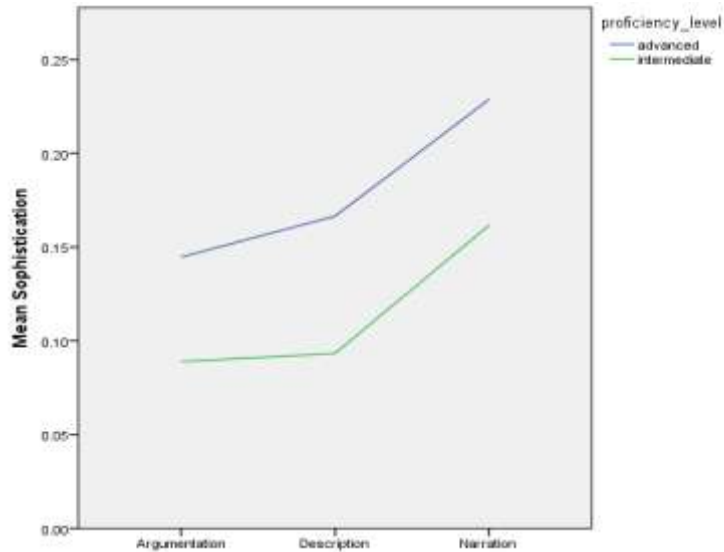


Figure 2. Intermediate and Advanced Learners' Mean Sophistication Scores across the Three Tasks

However, the interaction between task type and proficiency level was significant for density scores ($F(2, 66) = 6.146, p = .004, \text{effect size} = .157, \text{observed power} = .876$). The advanced learners did better than the intermediate learners on the description and narration tasks, but not on the argumentation task, as shown in Figure 3 below.

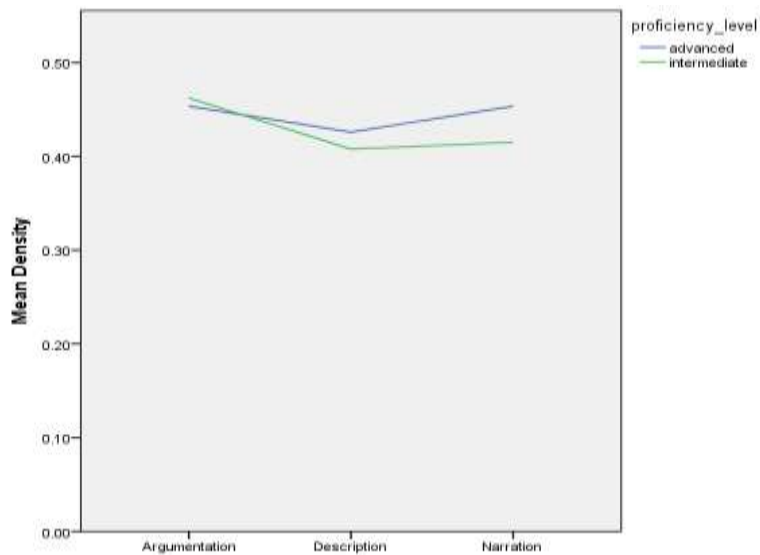


Figure 3. Intermediate and Advanced Learners' Mean Density Scores across the Three Tasks

4.1.2. The Effect of Proficiency Level on Lexical Complexity

The second research question addressed the effect of proficiency level on the participants' lexical performance. The results of the three ANOVA tests for the three different aspects of lexical complexity are shown in Table 4.

Table 4

Two-Way Repeated-Measures of ANOVA for Main Effect of Proficiency Level on Three Aspects of Lexical Complexity

	F	Sig.	Effect Size (η^2)	Observed Power
Diversity	12.291	.001	.271	.925
Density	5.213	.029	.136	.601
Sophistication	11.292	.002	.225	.903

As can be seen, the difference in the scores obtained by the advanced and the intermediate learners was significant for all the three different aspects of lexical complexity, but the effect sizes for diversity and sophistication were larger than the effect size for density.

4.2. Discussion

The results of the present study showed that there was a main effect of task type on the participants' lexical performance on the three speaking tasks. The argumentation and description tasks, as two non-narrative tasks, had significantly higher diversity scores than the narration task. The narration task, however, led to a significantly higher sophistication score than the other two non-narrative tasks. The significantly low diversity score for the narration task can be attributed to the fact that, in order to narrate a coherent story, the participants have to repeat part of what has already been said. Recycling of the words and phrases used in the previous sentences is needed to link the different episodes of the story together and create a coherent whole, hence a lower density score than the other two non-narrative tasks. The texts elicited by the argumentation and description tasks, however, can be divided into idea units that are independent from each other, eliminating the need to use lexical items already used in the previous sentences.

The effect of task type on the density aspect of lexical performance is not easy to explain. The description and narration tasks both result in density scores that are significantly lower than the argumentation task. One reason for the low density scores for these two tasks can be the frequent use of pronouns, hence increasing the number of function words in the text. In the description task, the participants were required to describe a friend they

enjoyed spending time with. The person was referred to by a third person singular pronoun after he or she was introduced in the beginning of the description. Likewise, the participants used the masculine third person pronoun to refer to the main character throughout the story. The significant difference between the description and narration tasks is difficult to explain.

As for the sophistication aspect of lexical complexity, the participants' differential performance across the three tasks can be attributed to the design features of the tasks. Narrative tasks are the most input-driven, and the least negotiable tasks which require the use of difficult-to-avoid lexical items implicated in the task (Skehan, 2009a). Narrative tasks make it difficult for the speaker to avoid using rare lexical items. The argumentation and description tasks are more negotiable and allow greater latitude. As Skehan (2009a) states, "Non-narrative tasks, although they make some lexical items salient, seem to allow participants freedom to express themselves without necessarily these key items if alternative means of expression can be found" (p. 121).

It was also found that there was a main effect of proficiency level. Advanced learners consistently did better than the intermediate learners on all the three aspects of lexical complexity across the three tasks. The only exception was the higher density score the intermediate learners had for density on the argumentation task. As proficiency increases, learners develop the ability to use a wider range of words and employ less frequent words while they are performing different task types. The reason for the higher density score obtained by the intermediate learners than the advanced learners on the argumentation task may be due to the fact that more proficient learners use more function words (e.g., subordinators and coordinators) to generate syntactically more complex utterances in argumentative tasks (Gass et al., 1999). Therefore, one needs to take the syntactic performance of learners into account to account for their variable lexical performance when it comes to lexical density. These results show that diversity and sophistication scores are able to discriminate the more proficient learners from the less proficient ones.

The results of the present study are in line with those obtained by Foster and Skehan (1996), Foster (2001), and Alexopoulou et al. (2017). Non-narrative tasks result in high diversity and low sophistication scores, while narrative task yield low diversity but high sophistication scores. One implication of this is that "the capacity to avoid recycling vocabulary, and the capacity to inject vocabulary richness into performance seem to connect with different aspects of speaking" (Skehan, 2009a, p. 120). Thus, we need to be specific about the generalizations we make about the effects of such independent variables as cognitive complexity and task type on L2 learners'

lexical performance, as the different aspects of lexical complexity (especially diversity and sophistication) behave differently on a specific task.

5. Conclusion and Implications

The findings of this study can shed light on the possibility of using these measures to distinguish more proficient from less proficient learners in automated tools of speech and writing assessment. Lexical diversity and lexical sophistication seem to be reliable measures to be used to distinguish speech produced by learners across different proficiency levels. One of the limitations of the present study is that it did not use data from native speaker performance as baseline data to see if the reported performances are the result of variable being manipulated in the study (task type in the present study) or simply the result of the second language speakerness of the learners (Skehan, 2009a).

References

- Ai, H., & Lu, X. (2010). A web-based system for automatic measurement of lexical complexity. *Paper presented at the 27th Annual Symposium of the Computer-Assisted Language Consortium (CALICO-10)*. Amherst, MA. June 8-12.
- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning, 67*(S1), 180-208.
- Brumfit, C. (1979). Communicative language teaching: An educational perspective. In C. J. Brumfit & K. Johnson (Eds.), *The communicative approach to language teaching* (p.183-191). Oxford: Oxford University Press.
- Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 23-46). Amsterdam/Philadelphia: John Benjamins.
- Daller, H., Van Hour, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics, 24*(2), 197-222.
- Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing, 28*(4), 561-580.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

- Foster, P. (2001). Rules and routines: a consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: second language learning, teaching, and testing* (pp. 75-93). Harlow: Longman.
- Foster, P. & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18(3), 299-324.
- Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, 30(1), 45-57.
- Gass, S., Mackey, A., Fernandez, M., & Alvarez-Torres, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 21(3), 354-375.
- Halliday, M. K. H. (1985). *An introduction to functional grammar* (1st ed.). London: Edward Arnold.
- Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency. Complexity, accuracy, and fluency in SLA*. Amsterdam/Philadelphia: John Benjamins.
- Jarvis, S. (2002). Short texts, best fitting curves, and new measures of lexical diversity. *Language Testing*, 19(1), 57-84.
- Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing*, 12(4), 377-403.
- Kuiken, F., & Velder, 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.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Lu, Xiaofei (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Language Journal*, 96(2), 190-208.
- McCarthy, P. M., & Jarvis, S. (2007). vocd: a theoretical and empirical evaluation. *Language Testing*, 24(4), 459-88.
- McCarthy, P. M. & Jarvis, S. (2010). MTLT, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Method*, 42(2), 381-392.

- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. [Computer Program]. <http://childes.psy.cmu.edu/>
- Malvern, D., & Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104.
- Malvern, D., Richards, B. J., Chipere, N, & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Basingstoke: Palgrave Macmillan.
- Meara, P., & Bell, H. (2001). P_lex: a simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(1), 5-19.
- Olinghouse, N.G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, 26(1), 45-65.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: CUP.
- Richards, B. J., & Malvern, D. D. (2007). Validity and threats to the validity of vocabulary measurement. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.). *Modelling and assessing vocabulary knowledge* (pp. 79-92). Cambridge: CUP.
- Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27-57.
- Robinson, P. (2003). The cognition Hypothesis, task-design and adult task-based language learning. *Second Language Studies*, 21(2), 45-107.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a Computational Framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, 43(1), 1-32.
- Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance*. Amsterdam/Philadelphia: John Benjamins.
- Robinson, P., & Gilabert, R. (2007). Introduction: Task Complexity, the Cognition Hypothesis, second language learning and performance. *International Review of Applied Linguistics in Language Teaching*, 45(3), 161-177.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1-14.

- Skehan, P. (2009a). Lexical performance by native and non-native speakers on language learning tasks. In B. Richards, H. M. Daller, D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application*. Palgrave Macmillan.
- Skehan, P. (2009b). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.
- Ure, J. (1971). Lexical density and register differentiation. In G. Perren and J. L. M. Trim (Eds.), *Applications of linguistics* (pp. 443-452). London: Cambridge University Press.
- Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239-236.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). Honolulu: University of Hawaii Press.
- Yoon, H.-J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275-301.

Bibliographic information of this paper for citing:

Bayazidi, A., Ansarin, A.-A., & Mohammadnia, Z. (2020). Lexical complexity as a function of task type and proficiency level in the speech monologs of Iranian EFL learners. *Journal of Modern Research in English Language Studies*, 7(1), 29-44.
