Applied Research on English Language

V. 10 N. 4 2021
pp: 39-70
http://jare.ui.ac.ir

DOI: 10.22108/ARE.2021.126660.1673
Document Type: Research Paper

The Effects of Task Type on Iranian EFL Learners' Use of Lexical Diversity and Sophistication

Ali-Akbar Ansarin [1*], **Mohammad Amin Karafkan** [2], **Yaser Hadidi** [3]

[1] *Professor, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran*
[2] *PhD Candidate, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran*
[3] *Assistant Professor, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran*

Received: 2020/12/26 **Accepted: 2021/05/03**

Abstract: EFL learners' ability to use advanced and varied vocabulary is a crucial issue in writing performance. This study aimed to investigate the effect of writing task type on lexical diversity and sophistication. The relationship between lexical sophistication and lexical diversity in the narrative, argumentative, and descriptive task types written by upper-intermediate EFL learners was also explored. To this end, 70 EFL upper-intermediate learners enrolled in an advanced writing course were selected as the participants. They undertook narrative, descriptive, and argumentative writing tasks in a counterbalanced way. Then, Vocabulary Diversity (Voc-D) and Measure of Textual Lexical Diversity (MTLD) were used to measure lexical diversity. Lexical sophistication was measured using CELEX log frequency and Beyond 2000 Words (K3+). The writings were pasted into Coh-Metrix measure MTLD, Voc-D, and CELEX. VocabProfiler was also used to measure the K3+ of the writings. In lieu of the purposes of the study, the research design involved repeated measures and correlation. Using Multivariate Analysis of Variance and Pearson's Correlation Coefficient, the findings revealed that lexical diversity varied significantly across the three writing tasks. The highest scores of MTLD and Voc-D belonged to the narrative, argumentative, and descriptive writing tasks, respectively. Moreover, EFL learners used more advanced words in argumentative task types according to the measurement of CELEX and K3+ level words. However, learners' lexical sophistication in narrative and descriptive writings did not emerge as significantly different when measured by K3+. The significant positive relationship between lexical sophistication and lexical diversity was evident in all three task types.

Keywords: Lexical Diversity, Lexical Richness, Lexical Sophistication, Writing Task Type.

* **Corresponding Author.**

Authors' Email Address:
[1] **Ali-Akbar Ansarin** (ansarin@tabrizu.ac.ir), [2] **Mohammad Amin Karafkan** (amin.karafkan2020@yahoo.com), **Yaser Hadidi** (hadidiy@tabrizu.ac.ir)

ISSN (Online): 2322-5343**, ISSN (Print):** 2252-0198

Introduction

Writing is considered a dynamic mind process by which the writer generates meaning. It is the most effective medium of conveying ideas and feelings via the text, yet a difficult skill for the majority of EFL learners to master (Dechant, 1993). Besides, as the primary conveyor of information load in communication, vocabulary plays a vital role in forming complete oral and written discourse, and sufficient vocabulary knowledge is a precondition for effective language use and comprehension of texts (Roche & Harrington, 2013). The writing process demands the writer to make careful choices of words and apply them simultaneously with other abilities to produce a text. Many studies have been done on learners' productive aspect of vocabulary knowledge based on the hypothesis, indicating that learners' vocabulary use may reveal their lexical knowledge and be strongly associated with their linguistic proficiency (Crossley, Salsbury, McNamara, & Jarvis, 2010).

In assessing writing proficiency or predicting future writing performance, assigning the writers only to one type of genre or task cannot bring out the language user's overall writing ability because each genre or discursive mode in writing demands different skills and knowledge (Yoon, 2018). Accordingly, it has been suggested that in order to assess L1 writing proficiency, it is necessary that three texts in each one of typical genres be scored by at least two expert raters (Bouwer, Beguin, Sanders & van den Bergh, 2015). Linguistic and discursive elements in L2 writing can affect judgments on writing quality (Crossley & McNamara, 2012; Kim & Crossley, 2018).

It has been established by genre researchers that the linguistic knowledge required for writing performance is different for different genres both in L1 and L2 (Biber, Gray, & Staples, 2016). Some researchers have therefore attempted to look for patterns of variation in the links between genre and writing performance in L1 and L2. There are numerous studies on L2 writing that examine the influence of task type on writing performance. However, these studies are mainly concerned with the effect of task type on overall scores of writing composition (e.g., Jeong, 2017; Qin & Uccelli, 2016) as well as syntactic aspects of writing task types (e.g., Casal & lee, 2019; Tavakoli, 2014), and there haven't been many studies looking at upper-intermediate Iranian EFL learners. This study highlights two aspects of lexical richness (i.e. lexical diversity and lexical sophistication) as they are widely used scoring standards for English proficiency exams such as TOEFL and IELTS. It investigates the sensitivity of lexical diversity and sophistication to

argumentative, descriptive, and narrative writing task types. Moreover, the relationship between lexical diversity and lexical sophistication has only been explored for a limited range of task types. The relationship between lexical diversity and lexical sophistication in argumentative, descriptive, and narrative task types will also be in focus. This study is significant in the steps it takes towards understanding upper-intermediate EFL students' productive knowledge and its role in their writing performance on different task types.

This study aimed at finding answers to the following research questions:

RQ1. Is there any significant difference in upper-intermediate EFL learners' narrative, descriptive and argumentative writing tasks in terms of lexical diversity?

RQ2.Is there any significant difference in upper-intermediate EFL learners' narrative, descriptive and argumentative writing tasks in terms of lexical sophistication?

RQ3. Is there any significant relationship between lexical diversity and lexical sophistication in narrative, descriptive, and argumentative writing tasks?

Literature Review

Lexical richness as a general measurement index for vocabulary quality (Milton, 2009) includes but is not limited to lexical diversity and lexical sophistication (Crossley et al., 2010). Lexical diversity and sophistication are used to measure lexical knowledge of L2 learners' oral and written language use. As a type/token ratio, lexical diversity is one of the measures used to describe the complexity of the written composition. It is the use of various unique words in the discourse. The writer or speaker ought to employ many unique words to develop a text with high lexical diversity (McCarthy & Jarvis, 2010). Earlier investigations (e.g., Crossley et al. 2010; McCarthy & Jarvis, 2010) suggest that lexical diversity as one of the most substantial indications of lexical competence is applied as a measure of evaluating linguistic performance. Furthermore, Santos (1988) revealed raters' judgments were mainly based on the correct or incorrect use of lexicon and the diversity of built-in lexical resources in the compositions. There are some methods including the number of types to tokens (TTR), Guiraud, Herdan, Uber, Maas, HD-D, Voc-D, and MTLD for measuring lexical diversity. Voc-D and MTLD are less sensitive to the text length (Bonvin & Lambelet, 2017). The Voc-D, designed by Malvern and Richards (1997), is described as an arithmetical transformation of the type-token ratio (TTR), which lessens text length influences. The Voc-D shows the extent of words' repetition in a discourse. MTLD, designed by McCarthy (2005), computes the mean length of consecutive word strings used in a text to maintain an above TTR value of .72 (McCarthy & Jarvis, 2010).

Lexical sophistication, also labeled as lexical rareness, is the relatively advanced or less frequent proportion of words in learners' writing (Read, 2000). The higher proficiency learners use rare words, regarded as sophisticated words, in their discourse, while less proficient ones use more frequent words that are less sophisticated (Crossley & McNamara, 2012; Kyle & Crossley, 2015; Zheng, 2016). Lexical sophistication is considered a multi-dimensional construct including *n*-gram frequency, hypernymy, polysemy, and psycholinguistic lexical information associated with word meaningfulness, imageability, concreteness, age of acquisition, and word familiarity (Kyle & Crossley, 2015). LFP, designed by Laufer and Nation (1995), can precisely determine the percentage of high-frequency words (first and second 1,000 words), academic words, and less frequent words used in a piece of writing. The words of the University Word List (UWL) and the not-in-the-lists categories were considered sophisticated or advanced, bearing in mind their rareness (Read, 2000) and low frequency of occurrence. CELEX, as a tool for measuring lexical sophistication, is a database of word frequencies adopted from a 17.9 million-word COBUILD corpus. Lower frequency scores belong to L2 learners employing less frequent words in their written discourse, and they have more advanced L2 vocabulary knowledge (Crossley et al., 2010). There are some factors such as age (Gharibi & Boers, 2017), educational context, and language proficiency (Kyle & Crossley, 2015; MacArthur et al. 2019), which may influence lexical richness. Some critical lexical-related factors, such as learners' receptive vocabulary breadth and depth, may directly affect their lexical richness in productive language use (Laufer & Nation, 1995).

Task type may induce learners to employ a particular set of lexical items in written discourse (Ruiz-Funes, 2015). Beauvais, Olive, and Passerault (2011) and Biber et al., (2016) argued that writing in different task types entails various cognitive and linguistic demands on behalf of the writers. A small body of investigations has also tried to examine the predictability of linguistic features of writing performance in multiple task types (e.g., Beers & Nagy, 2011; Yoon & Polio, 2016). Since there have been different ways of revealing ideas in each task type, some linguistic features may vary depending on task type (Beers & Nagy, 2011). According to Crowhurst and Piche (1979, as cited in Jagaiah, Olinghouse, & Kearns, 2020), descriptive and argumentative texts include a larger number of words per clause than narratives though Beers and Nagy (2011) came up with contradicting results by indicating more words per clause for narratives than descriptive and persuasive writings at least for higher-level learners.

Olinghouse and Wilson (2013) studied the role of vocabulary in the narrative, persuasive, and informative writings of EFL intermediate and upper-intermediate level learners. Their

study revealed that vocabulary diversity measured by MTLD was an exclusive predictor in narrative writing performance, whereas content words and registers were sole predictors for persuasive writing performance. Qin and Uccelli (2016) studied EFL Chinese secondary school learners' argumentative and narrative writing compositions and their scores for quality of writing and lexico-syntactic and genre-specific discourse features. They did not find any significant differences in quality ratings between the two task types. Moreover, they concluded that argumentative writings display a higher lexico-syntactic complexity. McNamara et al. (2013) also found out that length of sentences, word frequency measured by CELEX, word abstractness, the overlap of words across sentences, and the ratio of given to new information predicted first-year college learners' argumentative writing quality.

Woerfel and Yılmaz (2011) investigated learners' lexical diversity and text lengths in German-Turkish bilinguals' written discourses. They used Guiraud (G) and Advanced Guiraud (AG) as measures of lexical diversity. They found out that age, genre, and Turkish instruction influence lexical diversity and text length in language use. Moreover, Yoon and Polio (2016) investigated the role of linguistic features in argumentative and narrative essays written by ESL learners in the longitudinal study. They found out that argumentative written tasks were more associated with lexical sophistication, measured by CELEX and word length while narratives had greater lexical diversity measured by Voc-D. Durrant and Brenchley's (2019) investigated children's use of vocabulary in writing changes as they progress through their school careers. They investigated the extent to which a model of lexical sophistication as the use of low-frequency register-appropriate words adequately captures the development in vocabulary use across the course of compulsory education. They argued that lexical sophistication is inseparable from lexical diversity, a construct that is usually treated as distinct. Johnson, Acevedo, and Mercado (2016) argued that the use of high-frequency words was negatively correlated with L2 writing performance.

Vera Gomez et al. (2016) explored lexical knowledge in Chilean learners' narrative, persuasive, and expository writing skills. Using multilevel modeling, their study revealed that lexical diversity was more used in narrative and persuasive writing tasks, and lexical sophistication was a significant decisive feature for expository and narrative task types. 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 MTLD to measure global lexical complexity. The results showed that professional tasks resulted in the highest diversity scores, as compared with the description and

narrative tasks. Vogelin et al. (2019) investigated the influence of lexical features on teachers' judgments of English as a second language (ESL) argumentative essays. They found out that texts with greater lexical diversity and sophistication were assessed more positively concerning their overall quality as well as the analytic criteria 'grammar' and 'frame of the essay'. MacArthur, Jennings, and Philippakos (2019) explored the significant positive changes in linguistic features including word length (in syllables), word frequency, and age of acquisition in low-level learners' argumentative writing quality over the treatment time.

Juanggo (2018) studied lexical sophistication and diversity in B1 level and B2 level Indonesian EFL learners' argumentative written tasks. His study revealed a significant relationship between lexical diversity (measured by Voc-D) of their compositions and the use of the second most common 1000 wordlist. His study also indicated that there is not any significant difference between lexical diversity and sophistication of B1 level and B2 level Indonesian EFL learners. In contrast, Siskova's (2012) found a significant correlation between lexical diversity and sophistication in Czech EFL learners' narrative writings. The learners who use wider lexical variety can produce more sophisticated words in their narrative texts than students with a low lexical diversity index.

Some studies have been conducted in Iran to investigate the effects of task types on lexical richness. Sadeghi and Dilmaghani (2012) investigated task type effects on Iranian intermediate EFL learners' use of lexical diversity in writings. They found a significant difference between narrative and comparative, and also between narrative and argumentative genres in terms of lexical diversity (measured by Voc-D). Kalantari and Gholami (2017) found no significant correlation between lexical sophistication and lexical diversity in intermediate and advanced academic compositions. Bayazidi, Ansarin, and Mohammadnia (2019) investigated the correlation between different aspects of lexical richness in oral narrative, descriptive, and argumentative task types. They found out a positive correlation between lexical diversity (measured by Voc-D) and lexical sophistication (measured by beyond 2000 words), but the correlation was significant only for the oral narration task. Bayazidi, Ansarin, and Mohammadnia (2020) found the highest scores for lexical diversity (measured by Voc-D) in oral argumentation tasks, whereas the oral narrative task had the highest lexical sophistication measured by beyond 2000 words.

This investigation further explores the task type effect as the main factor in the productive vocabulary used by upper-intermediate Iranian EFL learners. In this regard, two lexical diversity and lexical sophistication measures will be used to explore their relationship with

each other in EFL learners' narrative, argumentative, and descriptive writing tasks

Methodology

Research Design

In order to respond to the first and second research questions, a repeated measures design was used to compare learners' written discourses in three different task types, as independent variables, with regard to the learners' scores of MTLD, Voc-D, K3+, and CELEX as four dependent variables. A correlational design was employed for the third question to explore the association between lexical diversity and lexical sophistication in three diverse task conditions including narrative, descriptive, and argumentative writing task types.

Participants

Seventy Iranian upper-intermediate EFL learners (49 males, 21 females) participated in this study from 6 advanced writing classes. Ranging from 20 to 26 years of age, the students were Persian-speaking undergraduate students studying non-English majors, all of them have been learning English for more than eight years at public schools, language institutes, and universities. The Oxford Quick Placement Test was employed to select those students who were at the upper-intermediate level. The nature of the writing course was to prepare the learners for IELTS, TOEFL, and Duolingo English exams based on a book called *Essay Becomes Easy*. Ryan's and Emma's writing videos were also used as supplementary materials. There were some reasons feeding into the choice of upper-intermediate EFL learners as the participants for the study. To begin with, based on Behjat and Sadighi's (2010) study, L1 to L2 positive transfer of writing skills cannot take place for learners with low-level linguistic proficiency. Moreover, as Ito (2009) found, this type of positive transfer can materialize for intermediate and advanced EFL learners. Another reason for selecting this proficiency level was the nature of the study, which involved the investigation of EFL learners' capacity to use their vocabulary knowledge in writing. The learners should have a good mental lexicon, the ability to use different words in context and, at the same time, a reasonably good command of English writing. The purposive sampling method was used to select learners familiar with descriptive, narrative, and argumentative writing task types.

Instruments

Oxford Quick Placement Test

The learners' level of proficiency was specified by the Oxford Quick Placement Test (OQPT) (2001, version 1) designed by Oxford University Press and the University of Cambridge Local Examinations Syndicate. This paper-based test consists of two parts: part one (Questions 1–40) and part two (Questions 41–60). The participants were supposed to answer this 60-item test, including vocabulary, grammar, and cloze tests in multiple-choice format, to be completed within 30-45 minutes. According to the scoring criteria of OQPT, scores ranging from 0 to 17 represent beginner levels, 18 to 29 elementary levels, 30 to 39 lower-intermediate levels, 40 to 47 upper-intermediate levels, 48 to 54 advanced levels, and 55 to 60 very advanced proficiency levels. As pertaining to the nature of this work, based on these criteria, learners scoring 40 to 47 were selected as upper-intermediate EFL learners in this study. Geranpayeh (2003) argued that OQPT, a standardized English proficiency test, has been pretested and validated by more than 6,000 students in about 20 countries. According to Allan (2004, as cited in Rashidi & Mirsalari, 2017), OQPT has been calibrated against the proficiency levels based on the Common European Framework of Reference for Languages (CEF), the Cambridge ESOL Examinations, and other major international examinations, such as TOEFL. This version of the test has been reported to have high validity and reliability, as reported by Hassaskhah and Hajikhani Roudsari (2015).

Writing Task Types

Three descriptive, narrative, and argumentative writing tasks were developed for this study (see Appendix A). The learners were asked to write a short essay in at least 250 words for each writing task type, a word size below 150 words is not considered reliable for measuring lexical frequency profile (Laufer & Nation, 1995). All writing tasks incorporated the familiar and related topic of success associated with learners' life issues. The topics were not too specialized to be easy for upper-intermediate learners to write about. The theme was adopted from Essay Becomes Easy, an established, well-endorsed main course book developed by equally established curriculum developers. The descriptive writing task asked learners to describe an individual they respect and admire most for his/her accomplishment. Some questions were provided to guide learners in the description of the related features. To elicit learners' narrative writing performance, they were instructed to narrate a memorable event to do with their accomplishments in life, including the context, actions, feelings, etc. The topic of the

argumentative writing tasks required learners to provide their reasons for their achievement in life.

Lexical Diversity Measure

The measure of vocabulary diversity (Voc-D) and measure of textual lexical diversity (MTLD), considered as the best indices for measuring lexical diversity values (McCarthy & Jarvis, 2010), were used in combination. The indices of MTLD and Voc-D were obtained from Coh-Metrix available on http://141.225.61.35/cohmetrix2017. The reliability and validity of Voc-D have been verified without the sample size problems found in earlier approaches (Crossley et al., 2010; Crossley et al., 2014; Malvern, Richards, Chipere & Durán, 2004; Yu, 2010). MTLD is also reliable and closely correlated with other lexical diversity indicators (McCarthy, 2005; Treffers-Daller & Korybski, 2015). Learners' writing tasks, following Yu's (2009) guidelines, were pasted into the Coh-Metrix to provide MTLD and Voc-D values.

Lexical Sophistication Measure

Lexical sophistication was operationalized in this study using two approaches: level-based and count-based vocabulary frequency. Lexical Frequency Profile (LFP) was used for the first approach. LFP, designed by Laufer and Nation (1995), can estimate the proportion of advanced vocabulary in the text. Beyond 2000 words, hereafter we call K3+, including academic and non-list words, were used in this study for measuring lexical sophistication in LFP. Proper nouns were eliminated when the learners' writings were pasted into LFP which is accessible at VocabProfiler (https://www.lextutor.ca/vp/eng/). Abbasian and Parizad (2011) verified the validity and reliability of LPF in measuring lexical richness in discourses of various types. Laufer and Nation (1995), and Nation (2006) argued that LFP is a reliable measure of level-based assessments.

Besides, we employ a count-based word frequency approach, which analyzes the frequency of individual words based on the corpus. It measures the mean frequency score for each text (Crossley et al., 2014). In this regard, CELEX word frequency (Baayen, Piepenbrock & Gulikers, 1995) (logarithm including content words), detected by Coh-Metrix, was used as another index of word sophistication in this study. The pasted writing tasks with omitted proper names were analyzed by Coh-Metrix to calculate the CELEX value. The reliability of this index was verified by Crossley, Cobb, and McNamara (2013).

Procedures

The informed consent form was distributed in the first session among the learners. The study's general objectives were briefly explained to the learners. They were informed that their performance on their writing tasks would not influence their final exam scores. We made sure that they took the written task seriously by respectfully explaining to them the seriousness of the tasks. Moreover, we respectfully reward them with a certain sum of money for the time they put in. This was a royalty extended to everyone who participated in this study. Then, in the first session, the Oxford Quick Placement Test was distributed among the learners to select the upper-intermediate participants. The participants completed the first part of the test within 30 minutes and finished the second part within 15 minutes.

From the 15th to 17th session of the writing course, the students were required to write the narrative, descriptive, and argumentative task types. Counterbalancing was adopted so as to mitigate the effects of cognitive loads and cognitive complexity. Then, 20 participants received narrative, descriptive, and argumentative tasks in the 15th, 16th, and 17th sessions, respectively. Another 20 participants received descriptive, argumentative, and narrative tasks in 15th, 16th, and 17th sessions, respectively. The rest of the participants (N=30) also received argumentative, narrative, and descriptive writing tasks in the following 15th, 16th, and 17th sessions, respectively. They wrote about success as the general topic using Notepad and a laptop and were asked to compose at least 250 words in 40 minutes about each given topic. Altogether, the data set consisted of 70 descriptive, 70 narratives, and 70 argumentative writing compositions. No dictionaries or any other reference tools were allowed while writing, and the researchers directly supervised the completion of the writing tasks.

Lexical diversity is aimed at investigating the variety of word ranges used, and the accurate and recognizable forms of the words must be used to achieve this central point. As the presence of errors in texts might have a confounding effect on lexical diversity ratings, the spelling mistakes in Notepad, pasted into Microsoft Word Office, were automatically corrected according to the following Yu's (2009) guidelines. First, several misspellings in the written texts were corrected and were not excluded from the analysis: one missing letter, one wrong additional letter, or one incorrectly placed letter. Also, other misspellings with more than one incorrect letter were regarded as unacceptable and were omitted. Additionally, the contracted forms such as *You're, didn't*, *doesn't*, etc., were detached. Finally, proper nouns, including the names of places or people in the writing tasks, were eliminated, and the topic prompts were

removed. Moreover, the texts were brought into the analyzers. In this study, the separate analysis of each learner's composition was done by VocabProfiler and Coh-Metrix to detect the K3+ and CELEX Log frequency for all words of each composition, respectively. Moreover, Coh-Metrix also provided Voc-D and MTLD as the indices of lexical diversity. For example, a learner's argumentative writing (see Appendix B) was checked for the removal of proper names and then was pasted into Coh-Metrix to measure its sophistication via CELEX Log frequency for all words (see Appendix C). Having used Yu's (2009) guidelines, the lexical diversity of the writing sample, measured by MTLD and Voc-D was obtained through Coh-Metrix (see Appendix D). The argumentative writing was also pasted into VocabProfiler to estimate the use of beyond 2000 words by adding academic and non-list words (see Appendix E).

Results

To select homogenous learners for this study, Oxford Quick Placement Test (OQPT) was administered to the initial 103 participants and, following the scoring guideline provided by the test, those whose scores fell within the range of 40 to 47 were selected as upper-intermediated learners. Table 1 shows the descriptive statistics of the initial and selected participants' OQPT scores.

Table 1. *Descriptive Statistics of the Initial and Selected Participants' OQPT Scores*

	N	Minimum	Maximum	Mean	Std. Deviation
Initial	103	36.00	53.00	44.0000	3.53415
Selected	70	40.00	47.00	43.0143	1.98905
Valid N (listwise)	70				

Results for Question 1

The first research question aimed to examine the difference in lexical diversity among the narrative, descriptive, and argumentative writings of EFL learners. Table 2 showed the descriptive statistics of the scores.

Table 2. *Descriptive Statistics of the Lexical Diversity Scores*

		N	Minimum	Maximum	Mean	SD	Skewness		
								Std.	
							Statistic	Error	Ratio
MTLD	Narrative	70	55.100	92.967	71.45233	9.989512	.339	.287	1.181
	Descriptive	70	41.556	77.781	58.27824	8.559191	.354	.287	1.233
	Argumentative	70	39.585	89.286	66.17537	11.914896	-.006	.287	-.021
Voc-D	Narrative	70	55.381	94.080	75.90377	8.642160	-.019	.287	-.066
	Descriptive	70	46.680	89.320	64.60511	9.960877	.438	.287	1.526
	Argumentative	70	46.641	94.261	70.92604	11.022076	.015	.287	.052
Valid N	(listwise)	70							

As reported in Table 2, the skewness ratios for all of the distributions were within the legitimate range of ±1.96, suggesting the normality of distributions.

The Multivariate Analysis of Variance (MANOVA) was run to answer this question. The lexical diversity of the writings was measured through two measures, that is, MTLD and Voc-D. Multivariate test (Table 3) demonstrated that there were significant differences between the three writing task types with regard to lexical diversity ($p = 0.000 < 0.01$, F = 13.999, Wilks' Lambda=.775).

Table 3. *Multivariate Test Results: Compare Overall Lexical Diversity in Three Task Types*

	Effect	Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.981	5417.418	2.000	206.000	.000
	Wilks' Lambda	.019	5417.418	2.000	206.000	.000
	Hotelling's Trace	52.596	5417.418	2.000	206.000	.000
	Roy's Largest Root	52.596	5417.418	2.000	206.000	.000
Writing Type	Pillai's Trace	.225	13.137	4.000	414.000	.000
	Wilks' Lambda	.775	13.999	4.000	412.000	.000
	Hotelling's Trace	.290	14.859	4.000	410.000	.000
	Roy's Largest Root	.289	29.880	2.000	207.000	.000

Table 4 demonstrates the test of between-subjects effects as part of the MANOVA output.

Table 4. *Tests of Between-Subjects Effects: Comparing MTLD and Voc-D in Three Task Types*

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	MTLD	6154.574[a]	2	3077.287	29.306	.000	.221
	Voc-D	4489.137[b]	2	2244.568	22.796	.000	.180
Intercept	MTLD	895513.230	1	895513.230	8528.296	.000	.976
	Voc-D	1043110.344	1	1043110.344	10593.819	.000	.981
Type	MTLD	6154.574	2	3077.287	29.306	.000	.221
	Voc-D	4489.137	2	2244.568	22.796	.000	.180
Error	MTLD	21736.024	207	105.005			
	Voc-D	20382.059	207	98.464			
Total	MTLD	923403.828	210				
	Voc-D	1067981.539	210				
Corrected Total	MTLD	27890.597	209				
	Voc-D	24871.195	209				

a. R Squared = .221 (Adjusted R Squared = .213)
b. R Squared = .180 (Adjusted R Squared = .173)

Table 4 indicated that three writing task types turned out to have a statistically significant differences both in MTLD ($F_{(2,207)} = 29.306$, $p = 0.000 < 0.01$) and Voc-D ($F_{(2,207)} = 22.796$, $p = 0.000 < 0.01$). Moreover, the effect sizes using Partial Eta Squared were 0.221 and 0.18 for MTLD and Voc-D scores, respectively, indicating that MTLD and Voc-D accounted for 22.1% and 18% of the overall variance of each of the corresponding dependent variables. Both of these values signified large effect sizes.

Finally, in order to locate the place of significant differences, a post hoc test was run (Table 5). The results showed that learners' narrative writings were significantly better than their descriptive writings in terms of lexical diversity when measured by MTLD (MD = 13.174, SE = 1.732, p = .000 < .01) and Voc-D (MD = 11.299, SE = 1.677, p = .000 < .01). However, when measured by MTLD (MD = 5.277, SE = 1.732, p = .007 < .01) and Voc-D (MD = 4.978, SE = 1.677, p = .009 < .01)), their scores of lexical diversity in their narrative writings were significantly higher than their lexical diversity scores in their argumentative ones. Finally, participants' scores of MTLD (MD = 7.897, SE = 1.732, p = .000 < .01), and Voc-D (MD = 6.321, SE = 1.677, p = .001 < .01) were significantly higher in argumentative writings compared to descriptive writings.

Table 5. *Multiple Comparisons: Tukey Post Hoc*

Dependent Variable	(I) Writing Type	(J) Writing Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
MTLD	Narrative	Descriptive	13.17409*	1.732092	.000	9.08514	17.26303
		Argumentative	5.27696*	1.732092	.007	1.18801	9.36590
	Descriptive	Narrative	-13.17409*	1.732092	.000	-17.26303	-9.08514
		Argumentative	-7.89713*	1.732092	.000	-11.98607	-3.80819
	Argumentative	Narrative	-5.27696*	1.732092	.007	-9.36590	-1.18801
		Descriptive	7.89713*	1.732092	.000	3.80819	11.98607
Voc-D	Narrative	Descriptive	11.29866*	1.677277	.000	7.33911	15.25820
		Argumentative	4.97773*	1.677277	.009	1.01819	8.93727
	Descriptive	Narrative	-11.29866*	1.677277	.000	-15.25820	-7.33911
		Argumentative	-6.32093*	1.677277	.001	-10.28047	-2.36139
	Argumentative	Narrative	-4.97773*	1.677277	.009	-8.93727	-1.01819
		Descriptive	6.32093*	1.677277	.001	2.36139	10.28047

Based on observed means.
The error term is Mean Square (Error) = 98.464.
*. The mean difference is significant at the .01 level.

Results for Question 2

This question aims to investigate the difference in lexical sophistication among EFL learners' narrative, descriptive, and argumentative writing tasks. The lexical sophistication of the writings was measured through two measures, that is, K3+ and CELEX. The descriptive statistics of the scores were shown in Table 6.

Table 6. . *Descriptive Statistics of the Lexical Sophistication Scores*

		N	Minimum	Maximum	Mean	SD	Skewness		
							Statistic	Std. Error	Ratio
K3+	Narrative	70	4.380	9.120	6.42614	1.249065	.340	.287	1.185
	Descriptive	70	3.960	8.490	6.04743	1.069704	.357	.287	1.244
	Argumentative	70	4.370	11.470	8.16771	1.701667	-.006	.287	-.021
CELEX	Narrative	70	2.628	4.031	3.21663	.324896	.315	.287	1.098
	Descriptive	70	2.527	4.428	3.42254	.447198	.098	.287	.341
	Argumentative	70	2.295	3.910	2.99023	.392704	.342	.287	1.192
Valid N (listwise)		70							

According to Table 6, the skewness ratios, within the legitimate range of ±1.96, suggest the normality of distributions.

Table 7 illustrates the Multivariate test: the value of Wilk's Lambda Test is 0. 537, F (2, 410) = 27.52 and $p = 0.000 < 0.01$, suggesting a statistically significant difference between the three task types concerning lexical sophistication.

Table 7. *Multivariate Test Results: Compare Overall Lexical Sophistication in Three Task Types*

	Effect	Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.003	39905.997[b]	2.000	206.000	.000
	Wilks' Lambda	387.437	39905.997[b]	2.000	206.000	.000
	Hotelling's Trace	387.437	39905.997[b]	2.000	206.000	.000
	Roy's Largest Root	.375	23.888	4.000	414.000	.000
Writing Type	Pillai's Trace	.641	25.699[b]	4.000	412.000	.000
	Wilks' Lambda	.537	27.521	4.000	410.000	.000
	Hotelling's Trace	.487	50.422[c]	2.000	207.000	.000
	Roy's Largest Root	.003	39905.997[b]	2.000	206.000	.000

In order to estimate the effect size, the test of between-subjects effects was run.

Table 8. *Tests of Between-Subjects Effects: Comparing CELEX and K3+ in Three Task Types*

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	CELEX	6.546[a]	2	3.273	21.358	.000	.171
	K3+	179.016[b]	2	89.508	47.950	.000	.317
Intercept	CELEX	2163.591	1	2163.591	14117.747	.000	.986
	K3+	9941.462	1	9941.462	5325.689	.000	.963
Type	CELEX	6.546	2	3.273	21.358	.000	.171
	K3+	179.016	2	89.508	47.950	.000	.317
Error	CELEX	31.723	207	.153			
	K3+	386.407	207	1.867			
Total	CELEX	2201.861	210				
	K3+	10506.885	210				
Corrected Total	CELEX	38.270	209				
	K3+	565.423	209				

a. R Squared = .171 (Adjusted R Squared = .163)
b. R Squared = .317 (Adjusted R Squared = .310)

Table 8 showed that CELEX ($F_{(2,207)}$ = 21.358, p = 0.000 < 0.01) and K3+ ($F_{(2,207)}$ = 47.95, p = 0.000 < 0.01) with effect size of 0.171 and 0.317, respectively, were significantly different in the three types of writing. CELEX and K3+ indicated large effect sizes as these measures accounted for 17.1% and 31.7% of the overall variance of each of the corresponding dependent variables.

To measure the significant difference between dependent variables in terms of CELEX and K3+, a post hoc test was run (Table 9). When measured by CELEX (MD = -0.226, SE = 0.661, p = .002 < .01), the learners' argumentative writings were significantly lower (better) than their narrative writings. However, their lexical sophistication in argumentative writings were significantly more than narrative ones when measured by K3+ (MD = 1.742, SE = 0.231, p = .000 < .01). The learners' argumentative writing tasks were significantly lower (better) than their descriptive writings in terms of CELEX (MD = -0.432, SE = 0.661, p = .000 < .01). Also, their use of lexical sophistication in argumentative writings, compared to descriptive ones, were higher when measured by K3+ (MD = 2.12, SE = 0.231, p = .000 < .01).

However, learners significantly outperformed in their narrative writings compared to their descriptive writings in terms of CELEX (MD = -0.206, SE = 0.661, p = .006 < .01) as

measure or lexical sophistication. In contrast, there was not any significant difference between participants' narrative and descriptive writings (MD = 0.379, SE = 0.231, p = .0231 > .01) in terms of lexical sophistication when measured by K3+.

Table 9. *Multiple Comparisons: Tukey Post Hoc to Locate Significant Differences Between Three Task Types*

Dependent Variable	(I) Writing Type	(J) Writing Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
CELEX	Narrative	Descriptive	-.20591*	.066171	.006	-.36213	-.04970
		Argumentative	.22640*	.066171	.002	.07019	.38261
	Descriptive	Narrative	.20591*	.066171	.006	.04970	.36213
		Argumentative	.43231*	.066171	.000	.27610	.58853
	Argumentative	Narrative	-.22640*	.066171	.002	-.38261	-.07019
		Descriptive	-.43231*	.066171	.000	-.58853	-.27610
K3+	Narrative	Descriptive	.3787	.23094	.231	-.1665	.9239
		Argumentative	1.7416*	.23094	.000	-2.2868	-1.1964
	Descriptive	Narrative	.3787	.23094	.231	-.9239	.1665
		Argumentative	2.1203*	.23094	.000	-2.6655	-1.5751
	Argumentative	Narrative	1.7416*	.23094	.000	1.1964	2.2868
		Descriptive	2.1203*	.23094	.000	1.5751	2.6655

Based on observed means.
The error term is Mean Square (Error) = 1.867.
*. The mean difference is significant at the .05 level.

Results for Question 3

The third question intended to examine the correlation between lexical sophistication and diversity in the narrative, descriptive, and argumentative writing tasks. The assumptions of the linear relation between each pair of variables, homoscedasticity, and normality of the distribution of variables were met to determine the legitimacy of running the analyses along with the parametric technique. Pearson's Correlation was used for the third research question.

According to Table 10, it was concluded that there was a positive and significant relationship between K3+, on one hand, and MTLD (ρ = .639, n = 70, p < .01) and Voc-D (ρ =

.491, $n = 70$, $p < .01$) in narrative writings on the hand in narrative writing, indicating that high levels of K3+ in narrative writing were associated with high levels of both measures of lexical diversity. Moreover, regarding descriptive writings, the correlation between K3+ and both measures of lexical diversity, that is, MTLD ($\rho = .575$, $n = 70$, $p < .01$) and Voc-D ($\rho = .504$, $n = 70$, $p < .01$) were significant and positive, as well. Finally, in the case of argumentative writings, significant and positive correlations were found between K3, on one hand, and MTLD ($\rho = .608$, $n = 70$, $p < .01$) and Voc-D ($\rho = .526$, $n = 70$, $p < .01$), on the other hand. Measured by Pearson's Correlation, the effect sizes between K3+ and MTLD in narrative, descriptive, and argumentative task types are large since the correlation is bigger than 0.5. However, there is a medium effect size between K3+ and Voc-D.

The results also showed that CELEX scores, for the case of narrative writings, have a significant and negative correlation with MTLD ($\rho = -.404$, $n = 70$, $p < .01$) and Voc-D ($\rho = -.502$, $n = 70$, $p < .01$), showing that low levels of CELEX in narrative writing were associated with high levels of both measures of lexical diversity.

Furthermore, as for descriptive writings, the relationships between CELEX and both measures of lexical diversity (i.e. MTLD ($\rho = -.455$, $n = 70$, $p < .01$) and Voc-D ($\rho = -.621$, $n = 70$, $p < .01$), were also significant and negative. As a final point, significant and negative correlations were found between CELEX, on one hand, and MTLD ($\rho = -.511$, $n = 70$, $p < .01$) and Voc-D ($\rho = -.635$, $n = 70$, $p < .01$), on the other in argumentative writings. There is a medium effect size between CELEX and MTLD in three writing tasks. Moreover, the effect size between CELEX and Voc-D in narrative, descriptive, and argumentative task types is moderately high.

Table 10. *Pearson's Correlation among Measures of Lexical Diversity and Lexical Sophistication*

		MTLD.N	MTLD.D	MTLD.A	Voc-D.N	Voc-D. D	Voc-D.A
K3+ .N	Pearson Correlation	.639**			.491**		
	Sig. (2-tailed)	.000			.000		
	N	70			70		
K3+. D	Pearson Correlation		.575**			.504**	
	Sig. (2-tailed)		.000			.000	
	N		70			70	
K3+. A	Pearson Correlation			.608**			.526**
	Sig. (2-tailed)			.000			.000

	N		70	70
CELEX.N	Pearson Correlation	-.404**	-.502**	
	Sig. (2-tailed)	.001	.000	
	N	70	70	
CELEX..D	Pearson Correlation	-.455**	-.621**	
	Sig. (2-tailed)	.000	.000	
	N	70	70	
CELEX.A	Pearson Correlation		-.511**	-.635**
	Sig. (2-tailed)		.000	.000
	N		70	70

Discussion

According to the first question, the study's findings demonstrated that lexical diversity as a component of lexical richness differs across narrative, descriptive, and argumentative writing tasks. The highest lexical diversity score, measured by Voc-D and MTLD, belonged to the narrative, argumentative, and descriptive writing tasks, respectively. Earlier studies have shown task type effects on linguistic features (Ruiz-Funes, 2015; Yoon & Polio, 2016). Our study results support Yoon and Polio's (2016) study, which indicated that narrative essays' lexical diversity was higher than argumentative ones. They found significant differences in the genre in the area of lexical diversity. Our results also tie well with Woerfel and Yılmaz's (2011) study, which showed the differences in lexical diversity among numerous genres. Moreover, our results align with Johansson's (2008) study, which found out that lexical diversity is higher in written narrative tasks than argumentative tasks. Our findings partially provide evidence to Olinghouse and Wilson (2013), who investigated the role of L1 vocabulary through L1 narrative, persuasive, and informative writing. Their study revealed that learners' use of lexical knowledge varies among L1 task types. Their lexical diversity and lexical maturity in narrative written discourse are more than informative and persuasive ones. They also found out that informative texts demonstrated the minimum amount of lexical diversity among the task type.

Our findings do not appear to corroborate Bayazidi et al.'s (2020) study. They found a significant low score in lexical diversity of the spontaneous speech in a narration task and the non-significant difference in the oral argumentation and description tasks' mean scores. In their study, participants have to replicate, recycle, and rephrase the words that they have already said to tell an intelligible and coherent narration in an oral form. They attributed their findings to Robinson's Cognition Hypothesis. According to Jackson and Suethanapornkul (2013),

Robinson's Cognition Hypothesis in explaining written task type effects on linguistic features is challenging since there are differences such as discourse mode and preparation time for planning written tasks. Another reason for the higher lexical diversity score in the narrative writing task in our study may be our learners' higher proficiency level.

Regarding the second question, our findings revealed that lexical sophistication in argumentative task types is significantly more than in narrative and descriptive tasks. In other words, EFL learners used more advanced words in argumentative tasks according to the measurement of CELEX and K3+ level words. However, learners' lexical sophistication in narrative and descriptive writings is not significantly different when measured by K3+. Our finding is consistent with Yoon and Polio (2016) in that longer words and less frequency level words are more remarkable in argumentative texts than narrative ones. They mentioned that nominalizations are widely used in argumentative tasks that may increase this task type's lexical sophistication. Our findings are also in line with Qin and Uccelli's (2016) study, which concluded that argumentative compositions showed a higher lexico-syntactic sophistication, measured using long words, abstract nouns, and academic words. In terms of lexical sophistication, our results were contrary to Vera et al.'s (2016) study, which demonstrated greater use of lexical sophistication in the expository and narrative texts than persuasive ones.

However, our results are contrary to Bayazidi et al.'s (2020) study, which found the oral narration task's highest sophistication score. They also attributed this finding to Skehan's (2009) perspective about various speaking dimensions in incorporating lexical richness into oral performance. Olinghouse and Wilson (2013) also found out that lexical maturity as a construct for determining lexical sophistication is higher in story text than argumentative text.

Regarding the third question, the lexical sophistication is significantly correlated with lexical diversity in the narrative, argumentative, and descriptive task types, and there are medium to large effect sizes between the lexical diversity and sophistication indices. In other words, those learners who use rare or infrequent words tend to activate a wider variety of words in three task types. This shows an interrelationship between learners' ability to use advanced words and their ability not to reuse the words or apply a wide variety of words in the narrative, argumentative, and descriptive tasks. Our findings are partially consistent with Bayazidi et al.'s (2020) study, which found the oral task type effect on the relationship between lexical diversity and sophistication. However, our findings are consistent with Bayazidi et al.'s (2019) study who found the large effect size between lexical sophistication and diversity in narrative discourse (r=.70). Moreover, our findings provide evidence to Siskova's (2012) study, which

revealed that advanced EFL learners who can use more advanced words in their written narrative discourse are more inclined to enjoy higher lexical diversity. Our findings contradicted Kalantari and Gholami's (2017) study, which reported that the lexical diversity and sophistication do not significantly correlate and there is a weak effect size between MTLD and beyond 2000 words (r=0.3). In contrary to our study, Juanggo (2018) examined lexical sophistication and diversity in B1 level and B2 level Indonesian EFL learners' written argumentative tasks. His study revealed that learners' use of the most varied vocabulary in argumentative writing is significantly correlated with the most common 2000 wordlist following the academic words and not in the list words. Apart from task type, there are some pedagogical and cognitive factors such as the learners' different learning context and learners' background knowledge which may influence their vocabulary diversity and lexical sophistication in their language use (Laufer & Nation, 1995; Siskova, 2012). Different learning contexts can influence L2 learners' lexical sophistication and diversity. Al Jahromi (2020) argued that the utilization of online discussions in L2 learning settings can enhance the use of lexical richness compared to traditional classrooms due to a number of interrelated cognitive factors. Lee and Anderson (2007) argued that the presence of topic familiarity positively influences productive skills in recalling the words and giving details on that topic.

Conclusion

This investigation revealed that lexical sophistication and diversity, as two constructs of lexical richness in productive language skills, should be considered by the practitioners, testers, and instructors. Raising learners' awareness about lexical diversity and sophistication in writing performance is of importance to instructors. It may help them reflect on their teaching methodology and the appropriateness of their teaching materials. Lexical richness should be a significant element to be taken into account in learner assessment and material development. Nonetheless, its implementation involves some difficulties. The learners' lexical diversity and sophistication in writing should be improved in EFL classrooms through multiple meaningful tasks to increase their vocabulary breadth for writing in different task types, particularly in narrative and argumentative ones. The learners' knowledge of semantics and word associations should also be developed either in classroom activities or in other curriculum materials. Students with more developed semantic networks are better in the vocabulary component of writing performance. A few empirical investigations have demonstrated that EFL learners'

word association knowledge could improve by input flooding and written and oral communication with other non-native and native speakers. Some supplemental vocabulary instruction based on word families can be presented to EFL learners to develop their lexical sophistication and increase their vocabulary knowledge in essay writing.

This study may motivate test designers to extend the scope of writing tasks in high-stakes tests. They can develop writing sections beyond the existing forms of writing ideal for high stakes examinations and come up with diverse writing task types and themes, which can activate diversity and improve thinking processes.

An apparent limitation here was the selection of a limited sample size. Only seventy male and female learners' lexical diversity and lexical sophistication were evaluated. Additionally, only one level of proficiency, that is, upper-intermediate level, was selected. Therefore, the findings clearly cannot be extrapolated to other proficiency levels. One suggestion is to investigate the lexical richness of other writing task types such as critical, deductive, and reflective types, comparing and contrasting them. Furthermore, this study used one general topic for writing purposes. Other studies can use different topics for each task type. Future researches are recommended to analyze larger samples with different proficiencies in the long term. Future work can also examine relationships between lexical complexities of ESL/EFL writing quality as judged by human raters in different task types to validate the assessment process. Future studies can look at relationships among writing topics and measures of lexical richness. Oral production, like writing, has been recognized as an intricate process since numerous features may interact to provide a holistic score eventually; therefore, it would be interesting to understand how lexical richness plays a role in L2 oral proficiency. Further studies can see how lexical richness develops in speaking performance.

References

Abbasian, G., & Parizad, M. (2011). Validation of lexical frequency profiles as a measure of lexical richness in written discourse. *Journal of Technology and Education, 5*(3), 221-227.

Al Jahromi, D. (2020). Can virtual noticing enhance L2 students' lexical complexity? A study of the effects of online discussions on L2 learners' lexical complexity. *Arab Journal of Applied Linguistics, 5*(2), 32-54.

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(1), 180–208.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bayazidi, A., Ansarin, A., & Mohammadnia, Z. (2019). The relationship between syntactic and lexical complexity in speech monologues of EFL learners. *Applied Research on English Language, 8*(4), 473-488.

Bayazidi, A., Ansarin, 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.

Beauvais, C., Olive, T., & Passerault, J. M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, *103*(2), 415–428.

Beers, S., & Nagy, W. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing, 24*,183–202.

Behjat, F., & Sadighi, F. (2010). Iranian English major students' L2 grammar development: Linguistic Threshold hypothesis. *The Electronic Journal for English as a Second Language, 13*(4), 6.

Biber, D, Gray, B & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels, *Applied Linguistics, 37*(5), 639-668.

Bonvin, A., & Lambelet, A. (2017). Algorithmic and subjective measures of lexical diversity in bilingual written corpora: a discussion. *Cognition, Representation and Language, 21*(1), 1-16.

Bouwer, R., Béguin, A., Sanders, T., & van Den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing, 32*(1), 83–100.

Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing, 44*(1), 51-62

Crossley, S.A., Salsbury, T., McNamara, D.S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing, 28*(4), 561-580.

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, *35*, 115–135.

Crossley, S. A., Cobb, T., & McNamara. D. S. (2013). Comparing count-based and band-based indices of word frequency: Implications for active vocabulary research and pedagogical applications, *System, 41*(1), 965-981

Crossley, S. A., Salsbury, T., & Mcnamara, D. S. (2014). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, *35* (1), 142-157.

Dechant, E. (1993). *Whole Language Reading: A Comprehensive Teaching Guide*. Lancaster, Pennsylvania: Technomic Publishing Company, Inc.

Durrant, P., & Brenchley, M. (2019). Development of vocabulary sophistication across genres in English children's writing. *Reading and Writing, 32,* 1927–1953.

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *Research Notes,* 12, 8–10.

Gharibi, K., & Boers, F. (2017). Influential factors in lexical richness of young heritage speakers' family language: Iranians in New Zealand. International Journal of Bilingualism, 1367006917728395.

Gómez, V. G., Sotomayor E, C., Bedwell, P., Domínguez, A., & Jéldrez, E. (2016). Analysis of lexical quality and its relation to writing quality for 4th grade, primary school students in Chile. *Reading and Writing, 29* (1), 1317–1336.

Hashemifardnia, A., Shafiee, S., Esfahani, F.,& Sepehri, M. (2021). Effects of Massive Open Online Course (MOOC) on Iranian EFL learners' speaking complexity, accuracy, and fluency. *Computer-Assisted Language Learning Electronic Journal, 22*(1), 56-79.

Hassaskhah, J., & Hajikhani Roudsari, F. (2015). The relationship between learners' academic self-concept and the consequential validity of Quick Placement Test. *International Journal of English and Education, 4*(4), 240-252.

Ito, F. (2009). Threshold to transfer writing skills from L1 to L2. *The bulletin of the Kanto-koshin-etsu English Language Education Society*, *23*, 1-10.

Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63*, 330-367.

Jagaiah, T., Olinghouse, N. G. & Kearns, D. M. (2020). Syntactic complexity measures: Variation by genre, grade-level, students' writing abilities, and writing quality. *Read Writ, 33,* 2577–2638.

Jeong, H. (2017). Narrative and expository genre effects on students, raters, and performance criteria. *Assessing Writing*, *31*, 113–125.

Johnson, M. D., Acevedo, A., & Mercado, L. (2016). Vocabulary knowledge and vocabulary use in L2 writing. *TESOL Journal, 7*(3), 700-715.

Juanggo, W. (2018). Investigating lexical diversity and lexical sophistication of productive vocabulary in the written discourse of Indonesian EFL learners. *Indonesian Journal of Applied Linguistics. 8*(1), 38-48.

Kalantari, R., & Gholami, J. (2017). Lexical complexity development from dynamic systems theory perspective: Lexical density, diversity, and sophistication. *International Journal of Instruction, 10*(4), 1-18.

Kim, M., & Crossley, S. A. (2018). Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing, 37*, 39-56.

Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *TESOL Quarterly*, *49*(4), 757-786.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307-322.

Lee, H., & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing, 24*(3), 307-330.

MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing, 32*(6), 1553–1574.

Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving Models of Language* (pp. 58–71). Clevedon: Multilingual Matters.

Malvern, D., Richards, J.B., Chipere, N., & Durán P. (2004). *Lexical Richness and Language Development: Quantification and Assessment.* Palgrave MacMillan.

McCarthy, P. M. (2005). *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)*. Unpublished PhD dissertation. University of Memphis.

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*, 381–392.

McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods, 45,* 499–515.

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*, 59–82.

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing: An Interdisciplinary Journal, 26*, 45–65.

Qin, W., & Uccelli, P. (2016). Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing, 33*, 3–17.

Rashidi, N., & Mirsalari, S. A. (2017). Investigating the relationship between Iranian EFL learners' use of strategies in collocating words and their proficiency level. *Journal of Research in Applied Linguistics, 8*(2), 93-118.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. *Language Testing in Asia, 3*(12), 1-13.

Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing, 28*, 1–19.

Sadeghi, K., & Dilmaghani, S. (2012). The relationship between lexical diversity and genre in Iranian EFL learners' writings. *Journal of Language Teaching and Research, 4*, 328-334.

Santos, T. (1988). Professors' reactions to the academic writing of nonnative-speaking students. *TESOL Quarterly, 22*, 69–90.

Siskova, Z. (2012). Lexical richness in EFL students' narratives. *Language Studies Working Papers, 4*, 26-36.

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchon (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 217–236). Amsterdam, Netherlands: John Benjamins.

Treffers-Daller, J., & Korybski, T. (2015). *Using lexical diversity measures to operationalize language dominance in bilinguals*. Retrieve from: https://www.researchgate.net/publication/294260607_Using_lexical_diversity_measures_to_operationalise_language_dominance_in_bilinguals.

Vera, G. G., Sotomayor, C., Bedwell, P., Domínguez, A. M., & Jéldrez, E. (2016). Analysis of lexical quality and its relation to writing quality for 4th grade, primary school students in Chile. *Reading and Writing*, *29*(7), 1317-1336.

Vogelin, C. Jansen, T., Keller, S. D., Machts, N., & Moller, J. (2019). The influence of lexical features on teacher judgments of ESL argumentative essays. *Assessing Writing, 39*, 50-63.

Woerfel, T. & Yilmaz, S. (2011), Lexical development of German-Turkish bilinguals: Acomparative study in written discourse In Chris Cummins et al. (eds.) *Proceedings ofthe 6th Cambridge Postgraduate Conference in Language Research*. Cambridge: Cambridge Institute of Language Research.

Yoon, H., & Polio, C. (2016). The linguistic development of students of English as a secondlanguage in two written genres. *TESOL Quarterly, 51*(2), 275–301.

Yoon, H. J. (2018). The development of ESL writing quality and lexical proficiency: suggestions for assessing writing achievement. *Language Assessment Quarterly, 15*(4), 387-405

Yu, G. (2009). Lexical diversity in writing and speaking task performances. *AppliedLinguistics*. *31*(2), 236-259.

Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System, 56*, 40-53.

Appendix A

Topics of Writing

Descriptive prompt

A person you admire most

There are many people around us with whom we live in different contexts such as home, workplace, school, university, etc. Some of them do not play any special role in our life, but some of them are so important. All of us know someone whom we admire a lot because s/he has some special characteristics or has done something for us. Describe the person you admire most for his/ her success in at least 250 words. You may use the following questions to know what you need to describe.

. Who this person is?

. How long you have known him/her?

. What qualities does he/she have?

. What does he look like? (Appearance)

. How is his/her personality?

. Why do you admire him/her?

Argumentative prompt

Some people believe that success in life comes from risks or chances. Others believe that success results from careful planning. In your opinion, what does success come from? Use specific reasons and examples to explain your position. You will have 40 min to complete this essay. A typical effective response would be at least 250 words.

Narrative prompt:

{A memorable event}

Write a personal story about a time when you achieved success. Please include detailed memories about that experience, including the context, your actions, feelings, etc. You will have 40 minutes to complete this narrative story. A typical effective response would be at least 250 words.

Appendix B

A sample of Argumentative Writing

To answer the question of this topic we can take different approaches. We can be dramatic and talk like motivational speakers about hard work and planning and being positive and things like this. But we all know that a lot of people try so hard and still never get to their goals at all. So if hard work is the reason it should be a reason of success for all people. So everyone who tries hard should achieve his or her goals. Again we know it is not a fact. I never mean that hard work is not important or we shouldn't have a plan or things like that. What I am trying to say is that there are other reasons that can have think about as reasons of success. Chance is surely one of the factors that successful people have and they always deny it. Because they think if they admit that they have chance it would mean they did not deserve the success they have achieved. But it is a truth that chance is important for success and it does not underestimate their success. Chance is there and successful people use it.

Another factor is being risk taker. People who are afraid of risks are not usually very successful and will end up staying in one place for ever with no possibility that they will succeed in future. To conclude we can say that success needs plan and hard work but chance and risk taking are also among the most important factors and shouldn't be ignored.

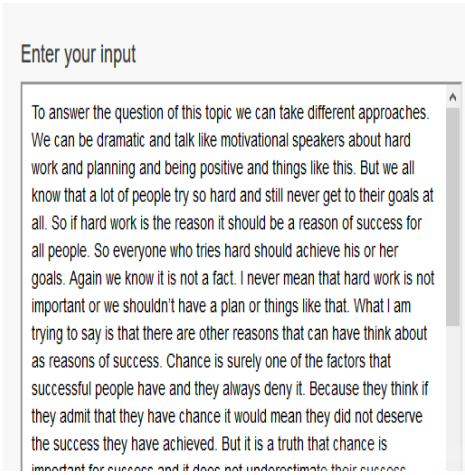
Appendix C

A sample of Calculating CELEX by Coh-Metrix, Using above Sample of Argumentative Writing

Word Information				
82	WRDNOUN	NOUNi	196.154	Noun incidence
83	WRDVERB	VERBi	111.538	Verb incidence
84	WRDADJ	ADJi	69.231	Adjective incidence
85	WRDADV	ADVi	80.769	Adverb incidence
86	WRDPRO	DENPRPi	100	Pronoun incidence
87	WRDPRP1s	n/a	7.692	First person singular pronoun incidence
88	WRDPRP1p	n/a	23.077	First person plural pronoun incidence
89	WRDPRP2	PRO2i	0	Second person pronoun incidence
90	WRDPRP3s	n/a	7.692	Third person singular pronoun incidence
91	WRDPRP3p	n/a	34.615	Third person plural pronoun incidence
92	WRDFRQc	FRCLacwm	2.746	CELEX word frequency for content words, mean
93	WRDFRQa	FRCLaewm	3.228	CELEX Log frequency for all words, mean
94	WRDFRQmc	FRCLmcsm	1.749	CELEX Log minimum frequency for content words, mean

Appendix D

A Sample of Calculating Voc-D and MTLD by Coh-Metrix, Using the Sample of Argumentative Writing



37	CRFCWOad	n/a	0.145	Content word overlap, all sentences, proportional, standard deviation
LSA				
38	LSASS1	LSAassa	0.235	LSA overlap, adjacent sentences, mean
39	LSASS1d	LSAassd	0.173	LSA overlap, adjacent sentences, standard deviation
40	LSASSp	LSApssa	0.196	LSA overlap, all sentences in paragraph, mean
41	LSASSpd	LSApssd	0.163	LSA overlap, all sentences in paragraph, standard deviation
42	LSAPP1	LSAppa	0.558	LSA overlap, adjacent paragraphs, mean
43	LSAPP1d	LSAppd	0	LSA overlap, adjacent paragraphs, standard deviation
44	LSAGN	LSAGN	0.330	LSA given/new, sentences, mean
45	LSAGNd	n/a	0.144	LSA given/new, sentences, standard deviation
Lexical Diversity				
46	LDTTRc	TYPTOKc	0.613	Lexical diversity, type-token ratio, content word lemmas
47	LDTTRa	n/a	0.462	Lexical diversity, type-token ratio, all words
48	LDMTLD	LEXDIVTD	62.466	Lexical diversity, MTLD, all words
49	LDVOCD	LEXDIVVD	75.102	Lexical diversity, VOCD, all words

Appendix E

A Sample of Calculating K3+ by VocabProfiler, Using above Sample of Argumentative Writing

	Families	Types	Tokens	Percent
K1 Words (1-1000):	84	98	**237**	**91.15%**
Function:	...	...	(145)	(55.77%)
Content:	...	...	(92)	(35.38%)
> Anglo-Sax	...	...	(49)	(18.85%)
K2 Words (1001-2000):	4	5	**6**	**2.31%**
> Anglo-Sax	...	...	(2)	(0.77%)
1k+2k			...	(93.46%)
AWL Words:	11	13	**15**	**5.77%**
> Anglo-Sax	...	...	(2)	(0.77%)
Off-List Words:	?	2	**2**	**0.77%**
	99+?	118	260	100%

Current profile	
%	Cumul.
91.15	91.15
2.31	93.46
5.77	99.23
0.77	100.00

Words in text (tokens):	260
Different words (types):	118
Type-token ratio:	0.45
Tokens per type:	2.20
Lex density (content words/total)	0.44
Pertaining to onlist only	
Tokens:	258
Types:	116
Families:	99
Tokens per family:	2.61
Types per family:	1.17
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%