



Depth and Breadth of Vocabulary Knowledge as Predictors of Narrative, Descriptive and Argumentative Writing

Mohammad Amin Karafkan^{1*}, Ali-Akbar Ansarin², Yaser Hadidi³

^{1*}Ph.D. Candidate, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran, amin.karafkan2020@yahoo.com

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

³Assistant Professor, Department of English, Faculty of Persian Literature and Foreign Languages, University of Tabriz, Tabriz, Iran, hadidiy@tabrizu.ac.ir

Article Info

ABSTRACT

Article type:

Research Article

Received:

19/10/2020

Accepted:

26/12/2020

Among intriguing areas in vocabulary acquisition research are such variables as breadth and depth of vocabulary knowledge, and their predictability in writing performance. In this spirit, this study set out to determine how receptive breadth, productive breadth, and receptive depth of word knowledge, using word families, predict total writing task score and the vocabulary component of EFL learners' narrative, descriptive, and argumentative writing performance. To this end, by administering Oxford Quick Placement Test to the learners enrolled in an advanced writing course, 70 (49 males, 21 females) EFL upper intermediate learners were selected as the participants of the study. To determine the participants' receptive depth, and productive and receptive breadth of the word knowledge, the Word Associates Test, the Lex30, and the Vocabulary Size Test were administered to the participants respectively. The participants also undertook descriptive, narrative and argumentative writing tasks. The results of the correlation coefficients and regression analyses of the data specified that: a) receptive vocabulary breadth and depth significantly contributed to both overall writing and vocabulary component of narrative, descriptive and argumentative writing; b) The breadth of productive vocabulary knowledge measured by the Lex30 only correlated with the vocabulary component score as well as the total score of narrative, descriptive, and argumentative writing. The implications include the fact that lexical knowledge aspects can be systematically used in both developing syllabus materials and classroom teaching methodologies.

Keywords: L2 Argumentative Writing, L2 Descriptive Writing, L2 Narrative Writing, Vocabulary Breadth, Vocabulary Depth

Cite this article: Karafkan, M. A., Ansarin, A. A., & Hadidi, Y. (2022). Depth and breadth of vocabulary knowledge as predictors of narrative, descriptive and argumentative writing. *Journal of Modern Research in English Language Studies*, 9(2), 27-50. DOI: 10.30479/jmrels.2020.14268.1757



© The Author(s).

Publisher: Imam Khomeini International University

1. Introduction

Learning vocabulary, as the primary conveyor of information load in communication, has an important role in developing complete oral and written discourse, and sufficient vocabulary knowledge is a precondition for effective language use and comprehension of texts (Roche & Harrington, 2013). In order to define knowledge of a lexical item, in recent decades, multiple complementary frameworks have been developed, which assume the multiple aspects of the vocabulary knowledge (Read, 2000). A well-known dimension of word knowledge may be the distinction between vocabulary breadth and vocabulary depth (Nation, 2001).

As the main constituent of a learner's lexical proficiency, vocabulary breadth is mostly used in applied linguistics and language teaching literature. It refers to the quantity of the words the learner knows or has partial meaning knowledge at a particular time (Nation, 2001). Vocabulary depth, instead, is regarded as another main facet of lexical competence. Vocabulary acquisition researchers agreed that vocabulary depth is related to the quality of the learner's knowledge about a lexical item (Schmitt, 2014). Nation (2001) also pointed out that the vocabulary depth involves not only semantic knowledge but also a wide variety of other dimensions of knowing a word, including paradigmatic (antonymy, synonymy, hyponymy) and syntagmatic characteristics (collocational). Schmitt (2014) investigated the construct of vocabulary depth concerning the word associations. He argued that the associative approach to vocabulary depth was to measure the learner's ability to find the relationship between the lexical item and other related words.

Albrechtsen et al. (2008) and Lee (2003) defined vocabulary knowledge as the vocabulary breadth and found its primary contributor to writing performance. They did not mention the vocabulary depth as an aspect of word knowledge. This is also evident in a few studies where the intent was to investigate explicit vocabulary instruction's influence on improving writing skill (Muncie, 2002) and productive use of vocabulary in an immediate writing task (Lee, 2003). The multidimensionality of word knowledge implies that the growth of vocabulary size is not sufficient to create a rich vocabulary knowledge repertoire (Nation, 2001). The focus of the investigations on the knowledge of vocabulary depth was mainly on its role in receptive skills like reading comprehension (e.g., Akbarian, 2010; Alavi & Akbarian, 2012; Cheng & Matthews, 2016; Mehrpour et al., 2011), in listening comprehension (e.g., Matthews, 2018; Teng, 2016), in L2 learners' use of lexical inferencing and success (Nassaji, 2006), and its relationship with grammatical knowledge (Kaivanpanah & Zandi, 2009). The apparent significance of vocabulary used in writing as a productive skill persuaded the researchers to study different aspects of word knowledge and their effect on writing performance. Also, very few investigations have

considered the contribution of vocabulary dimensions to a particular type of writing task, such as descriptive writing (Dabbagh & Janebi Enayat, 2017). Moreover, one of the leading causes of poor writing is the students' lack of enough lexical knowledge. In effect, foreign or second language learners regularly consider insufficient word knowledge as a primary barrier in the receptive and productive use of language (Richards, 2008). Therefore, language teachers should be conscious of learners' word knowledge dimensions and their use in writing tasks

To the best of our knowledge, the role of breadth and depth of vocabulary knowledge in predicting writing performance in different types of writing tasks seems not to be extensively investigated, which makes for an impetus to do this study. This study inspects the predictability of upper-intermediate EFL learners' receptive knowledge of vocabulary depth versus receptive and productive knowledge of vocabulary breadth in writing performance. It aims at determining whether productive and receptive knowledge of vocabulary breadth and receptive knowledge of vocabulary depth predict EFL learners' overall writing performance as well as in vocabulary component of descriptive, argumentative and narrative task types.

2. Literature Review

2.1. The Related Studies on the Relationship Between Vocabulary Breadth, Depth, and Writing

It is assumed that word knowledge has critical roles in productive skills such as writing, since words convey the information load of the meanings to the readers to comprehend (Olinghouse & Wilson, 2013; Schoonen et al., 2011). Miralpeix and Muñoz (2018) investigated the correlation between EFL higher-level learners' receptive breadth of word knowledge and their four language skills and their study showed that their receptive breadth of word knowledge is highly correlated with writing performance and is moderately related to reading, speaking and listening. Stæhr (2008) also explored the contribution of EFL Danish lower-level learners' receptive vocabulary size to their English listening, reading, and writing skills, and his study indicated that receptive vocabulary breadth was moderately correlated with their listening skill and closely related to writing and reading skills.

In fact, vocabulary knowledge as a benchmark of language proficiency may influence learners' quality of writing (Milton, 2013). Word knowledge works as a vital tool for L2 learners to develop written discourses for practical communicative purposes (Azodi et al., 2014). Therefore, vocabulary-based instruction using different vocabulary learning strategies is likely to develop writing performance (Levitzky-Aviad & Laufer, 2013).

Further, the significant relationship between receptive vocabulary sized and writing skill was also proved by Llach and Gallego's (2009) and Albrechtsen, et al.'s (2008) studies who asserted that learners with higher scores on receptive vocabulary size tests commonly receive higher scores in writing composition. Using the structural equation model, Schoonen et al., (2011), however, found out that EFL learners' receptive vocabulary knowledge was not a significant contributor to L2 writing proficiency. However, they found out that EFL writing was moderately related to lexical knowledge.

Li and Kirby (2014) operationalized the knowledge of vocabulary depth as the provision of the exact word, choosing the correct usage of words and correct identification of morphological word formation. They found that vocabulary depth was closely and meaningfully correlated with summary writing. They argued that quality of word knowledge is useful in eliciting denotation of words from the text and in writing a summary with some evidence for key notions. In contrast, Silverman et al. (2015) pointed out that productive depth of vocabulary knowledge and productive breadth of word knowledge, in comparison with the knowledge of receptive vocabulary depth, make a more evident contribution to the content, word choice, and the style of narrative writing tasks.

The productive breadth of the word knowledge and its relevance to productive skills have also attracted interests from investigators (Kilic, 2019; Laufer & Nation, 1995). A recent study by Kilic (2019) and Johnson et al. (2016) indicated that productive word knowledge measured by Productive Vocabulary Level Test (PVLTL) was moderately associated with the word choice in L2 writing performance. Johnson, et al. (2016) mentioned that learners' productive low-frequency words are positively associated with their overall writing performance. Regarding the relationship between the performance on the Lex30 as a productive test for the breadth of word knowledge and linguistic skills, Uchiyama and Saito (2019) explored the relationship between Japanese learners' performance on the Lex30, as a productive vocabulary knowledge test, and their speaking skill, and they found that the productive word knowledge scores predicted L2 fluency, but not comprehensibility. However, to date, the correlation between the performance on the Lex30 test and writing as a productive skill remained unknown.

Empirical investigations have supported the relationship between productive word knowledge and writing quality by measuring lexical diversity and lexical sophistication in free language production (Jeong, 2017). Johnson et al. (2013) found a significant relationship between learners' lexical sophistication, measured by the lexical frequency profile, and their holistic scores in essays. Engber (1995) studied the influence of ESL adult students' word knowledge on L2 writing performance and found

out that those who employ a great number of various words, receive higher scores in their writing based on a holistic rubric. In line with Engber (1995), Bulté and Housen (2014) investigated the quality of written texts using an analytic rubric, and they indicated that the correlation between vocabulary scores and lexical diversity was positive and significant. Therefore, raters should be instructed to assess the essays based on the lexical diversity, lexical accuracy and lexical sophistication which may improve the construct validity of the test for measuring productive lexical knowledge (Fritz & Ruegg, 2013).

Beauvais et al. (2011) argued that writing the different task types entails various cognitive task loads. Biber et al. (2016) also mentioned that task types require writers' linguistic knowledge. A small body of investigations have tried to examine and compare the predictability of lexical features in writing quality of argumentative, narrative, descriptive, expository task types. (e.g., Juanggo, 2018; Ruiz-Funes, 2015; Yoon & Polio, 2017).

Dabbagh and Janebi Enayat (2017) investigated the role of breadth and depth of word knowledge in descriptive writing task as one of the different task types. They used VLT and WAT for measuring the receptive vocabulary size and depth, respectively. Their study revealed the prediction of vocabulary size in the overall score of L2 descriptive writing performance. However, they concluded that vocabulary depth, only correlated with L2 learners' writing performance. Studies investigating the role of the two vocabulary dimensions (breadth and depth) in different writing task types are rare. Moreover, no study has explored the relationship between EFL learners' receptive breadth of word knowledge measured by VST and their writing performance. So, this study aimed at investigating the role of EFL learners' word knowledge (receptive and productive aspects of breadth and receptive aspect of vocabulary depth) in predicting the overall writing performance of different task types including descriptive, narrative and argumentative writings as well as in the vocabulary component of these task types

The present study tries to answer these questions:

1. Do the scores of the VST, Lex30, and WAT predict the overall assessment scores of upper-intermediate EFL learners' narrative, descriptive, and argumentative writing tasks?
2. Do the scores of the VST, Lex30, and WAT predict the vocabulary component scores of the upper-intermediate EFL learners' narrative, descriptive, and argumentative writing tasks?

3. Method

3.1. Design of the Study

This study used a descriptive correlational design. In this quantitative study, learners' scores on the Vocabulary Size Test (VST), Lex30, and Word Associates Test (WAT) acted as predictor variables. Their overall writing scores and vocabulary component scores in the narrative, descriptive, and argumentative writing tasks were regarded as criterion variables. This study's nature of variables is interval as they have a numerical value, and they can be gauged along a continuum.

3.2. Participant

103 EFL learners enrolled in the advanced writing course in the adult departments at four private language institutes, and took part in the study. The Oxford Quick Placement Test (2001, version1) was applied to select those students who were at the upper-intermediate language proficiency level. Seventy upper-intermediate learners (49 males, 21 females) were selected for the final analysis of the study from six advanced writing classes at four English language institutes. They were undergraduate students studying non-English majors from 20 to 26 years old, and all had been learning English for more than eight years at public schools, language institutes and universities.

Two qualified EFL instructors also contributed to this study as evaluators of the narrative, descriptive, and argumentative writing tasks. They were PhD candidates in English Language Teaching who had obtained a score of 8.5 on the IELTS writing section. They had upwards of five years of teaching experience under their belt, direct experience with teaching writing, in language schools and universities in Iran.

3.3. Instruments

3.3.1. *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 University of Cambridge Local Examinations Syndicate. This 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, which includes vocabulary, grammar, and cloze tests, in the 30-minute allocated time. According to the scoring criteria of OQPT, scores ranging from 0 to 17 represent the beginner level, scores ranging from 18 to 29 show elementary level, scores from 30 to 39 indicate lower-intermediate level, and scores between 40 and 47 are considered to be at an upper-intermediate level. Moreover, scores ranging from 48 to 54 and scores between 55 and 60 are

taken to be at advanced and very advanced proficiency levels, respectively. Based on these criteria, learners who scored 40 to 47 were selected as upper-intermediate EFL learners in this study.

3.3.2. Vocabulary Size Test

Vocabulary Size Test (VST), version A, designed by Nation and Beglar's (2007), was used to test the breadth of the receptive word knowledge. It is a non-diagnostic proficiency test intended to estimate EFL learners' understanding of vocabulary and measure their partial word knowledge. A large number of investigations support the use of VST and endorse its reliability and validity (Elgort, 2013; Karami, 2012). For instance, Beglar (2010) called VST the standardized test and proved its high reliability by Rasch reliability indices, greater than 0.96. VST consists of fourteen levels of word frequency, which begins with 1000 words and ends with 14000 most frequent English word families, derived from the fourteen 1000 BNC word lists (Nation & Beglar, 2007). Each level consists of 10 items and 40 choices. The VST is freely available online at www.lextutor.ca.

3.3.3. Word Associates Task

Word Associates Task (WAT) was originally designed by Read (1993) to estimate the receptive aspect of vocabulary depth knowledge. It measures numerous collocational and semantic relationships that a word has with other lexical items in a language. Qian (1998) later called this test the Word Associates Test (WAT). The obtained reliability for this test has reached 0.93 and 0.91 with a sample of 74 Korean and 94 Chinese native learners, respectively (Qian, 1998). The test covers 40 items, each one providing an adjective stimulus word above two boxes including four words. Most of the prompt words are general academic adjectives. Among the four words in the left box, one to three words can be synonymous with one aspect of the adjective or whole facets of its meaning. One to three words, among the four words in the right box, may have collocation with the stimulus word. Learners received one point for each correct answer. Test takers were not penalized for wrong responses. The maximum score is 160 in this test.

3.3.4. Lex30

Lex30 was developed by Meara and Fitzpatrick (2000) to estimate learners' breadth of the productive vocabulary knowledge, and it measures one component of productive vocabulary knowledge called word recall (Schmitt, 2014). It is proved by some strong evidence that Lex30 has high validity and reliability to measure the productive vocabulary knowledge (Uchihara & Saito, 2019; Walters, 2012). This test contains 30 stimulate words for which non-native speakers are assumed to write up to only four English words that they recall (Fitzpatrick & Meara, 2004).

3.3.5. Writing Task Types

Three descriptive, narrative, and argumentative writing tasks were developed for this study. The learners were asked to write a short essay in at least 350 words for each writing task type. 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 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 of their accomplishment in their life. The topic of argumentative writing tasks required learners to provide their reasons for their achievement in life.

The writing tasks were assessed using a modified analytic scale developed by Polio and Lim (2020). This analytic scale can assess the descriptive writing task (Dabbagh & Janebi Enayat., 2017) and narrative and argumentative writing tasks (Yoon, 2018). This analytic scale can supply more comprehensive information about learners' performance in various writing features and thus can offer a more consistent assessment scale (Yoon, 2018). Connor-Linton and Polio (2014) and Yoon (2018) confirmed the high reliability and validity of this scale in evaluating writing tasks. Dabbagh and Janebi (2017) reported a high inter-rater reliability coefficient for the total writing score ($r = .96$). The inter-rater reliability, obtained by Yoon (2018), showed acceptable reliability for this analytic scale ($r = .87$). In this study, two assessors reached an acceptable inter-rater reliability as measured by the Pearson's correlation for narrative task type ($r = .814$), argumentative task type ($r = .843$), and for descriptive task type ($r = .864$).

3.4. Procedure

The learners who enrolled in an advanced writing course for one and half hours in eighteen sessions during two months participated in this study. The study took six sessions. During the 18 sessions of the writing course, the students of four classes at two language institutes were instructed based on a book called *Essay Becomes Easy*. Ryan's and Emma's writing videos were also used as supplementary materials. The informed consent form was distributed in the first session among the learners. The study's general objectives were briefly explained to the learners, and they were informed that their performance on the vocabulary tests and writing tasks would not have any influence on 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 every one

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 test was administrated for 30 minutes. In session 6, WAT and the pencil-and-paper format of Lex30 were applied. They were informed about Lex30 and WAT. The time allocated to each of these tests was 30 minutes. In session 8, VST was distributed among the learners. Learners were instructed not to respond to the items they did not know to lessen the influence of guessing. The learners completed the test in 40 minutes.

From the 15th to 17th session of the writing course, the students were required to write the narrative, descriptive, and argumentative task types, respectively. These respective writing tasks were sequenced based on the syllabus of writing course and the cognitive loads and cognitive complexity. They wrote about the general topic of *Success* by the Notepad Program using a laptop since scribble, and illegible handwriting or various spelling errors, as surface-level features, could negatively affect the scores of the writing assessment (Graham et al., 2011). They were required to take 5 minutes to plan before writing, and then they were required to write at least 350 words in 40 minutes about each 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 giving vocabulary exams and the researcher directly supervised the completion of the vocabulary tests, and writing tasks.

To score learners' performance on Lex30, the answers were lemmatized based on Bauer and Nation's (1993) guidelines. Then the learners' written words were typed into the following online text analyzer website called VocabProfile: <http://lex tutor.ca>. The Vocabprofile scored the words based on different vocabulary levels, including the first 1000 words (K1), the second 1000 words (K2), the Academic Word List (AWL), and the off-list words. According to Meara and Fitzpatrick's (2000) scoring approach, 0 point was given not only for K1 as the most frequent 1000 words, but also for missing, unreadable, or non-existing English words, proper names, and numbers. Other responses beyond the first 1000 words scored one point.

3.5. Data Analysis

In this descriptive correlational study, the Kolmogorov-Smirnov test was employed to specify the data normality. Since the data were not normally distributed, Spearman correlation coefficient tests were run to examine the correlation between the scores on VST, WAT, Lex30 and overall score and vocabulary component score of narrative, descriptive and argumentative

writing tasks. This research project enjoyed multiple linear regressions using SPSS version 20.

4. Results and Discussion

4.1. Results

Descriptive statistics of learners' performance on vocabulary tests and writing tasks are presented in Table 1. This study used six regression models in order to answer the first and second questions (see Table 4 and Table 7).

Table 1

Descriptive Statistics for Learners' Scores on Predictor and Criterion Variables

		Minimum	Maximum	Mean	Std. Deviation
VST	70	59.00	111.00	87.68	11.85
WAT	70	72.00	128.00	96.51	15.87
LEX30	70	24.00	73.00	41.62	14.06
Narrative. T	70	37.00	77.00	57.18	8.96
Descriptive. T	70	37.00	81.00	58.75	9.75
Argumentative. T	70	31.00	71.00	54.14	8.75
Narrative. V	70	7.00	19.00	13.69	2.62
Descriptive. V	70	9.00	20.00	14.37	2.67
Argumentative. V	70	7.00	17.00	12.33	2.35
Valid N (listwise)	70				

Before analyzing the regression models, some preliminary assumptions should be inspected. In Table 2, Kolmogorov-Smirnov test was run in order to systematically check the normality of the distributions. As presented in Table 2, the Sig. values for the scores of VST (.05), WAT (.01), and Lex30 (.00) are not above the critical value (.05). Therefore, the normality of distribution for none of the scores is supported (Tabachnick & Fidell, 2013) and the data did not perfectly meet the presumptions of parametric statistical technique.

Table 2

Tests of Normality for the VST, WAT, and Lex30 Scores

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
VST	.10	70	.05	.94	70	.00
WAT	.11	70	.01	.94	70	.00
LEX30	.18	70	.00	.89	70	.00

a. Lilliefors Significance Correction

The legitimacy of exploring the predictability of the predicting variables depends on the presence of significant correlations between each of the predicting variables and the correspondent criterion variables in the regression models. Since the scores on the vocabulary tests lacked normal distribution, a non-parametric test like Spearman's Rank-Order correlation coefficient was used for examining the pre-requisite of the first and second questions.

As Table 3 shows, there are significant correlations between each pair of the predictor and criterion variables. Therefore, the pre-requisite for running the regression analyses is met.

Table 3

Spearman's Rank-Order Correlation among Predictor and Criterion Variables in the Regression Models

			VST	WAT	Lex30
Spearman's rho	Narrative. T	Correlation Coefficient	.55**	.59**	.35**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70
	Descriptive. T	Correlation Coefficient	.59**	.56**	.33**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70
	Argumentative. T	Correlation Coefficient	.54**	.56**	.35**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70
	Narrative. V	Correlation Coefficient	.70**	.68**	.40**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70
	Descriptive. V	Correlation Coefficient	.71**	.62**	.35**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70
	Argumentative. V	Correlation Coefficient	.65**	.69**	.34**
		Sig. (2-tailed)	.00	.00	.00
		N	70	70	70

** . Correlation is significant at the 0.01 level (2-tailed).

However, other assumptions, including multicollinearity, regression normality, and outliers, have to be examined before doing the analysis. The Tolerance Value and VIF were measured to confirm or reject the existence of multicollinearity. Tolerance values were more than .1, and the VIF values were below 10. Therefore, it concludes that multicollinearity did not occur for all the same predictors in six models.

The Normal Probability Plot (P-P) of regression standardized residuals and the scatterplot of standardized residuals for the six regression models were visually examined to test the normality in a regression analysis. Normal P-P showed no major deviation from normality. Furthermore,

scatterplot of standardized residuals showed that residuals were rectangularly dispersed, with most of the scores concentrated in the center.

In order to inspect the presence of outliers for the six models, Mahalanobis Distance values were also examined. The highest Mahalanobis value was 8.665, which was safely smaller than the critical level, suggesting the absence of an outlier case. Cook's Distance values were appropriately below the critical value 1. Consequently, assumptions including multicollinearity, regression normality, and outliers were not violated, and it was reasonable to run a multiple regression to respond to the first and second research questions.

4.1.1. The Contribution of Scores on VST, WAT and Lex30 to the Overall Scores of Narrative, Descriptive, and Argumentative Writing

In order to answer the first question, three standard multiple linear regressions were run on the writing scores of narrative, descriptive, and argumentative writing tasks. Table 3 presents the summary for regression models, including the R and R^2 .

Based on Table 4, R square came out to be 0.47, 0.45 and 0.43 for the first, second and the third model respectively. In other words, the first, second and third model indicate 47.3, 45.1 and 43.3 percent of the variance in total score of the narrative, descriptive and argumentative writings respectively.

Table 4

Regression Model Summary – R and R Square

Model ^a	R	R Square	Adjusted R Square	Std. Error of the Estimate
1 ^b	.68	.47	.44	6.65
2 ^c	.67	.45	.42	7.38
3 ^d	.65	.43	.40	6.74

a. Predictors: (Constant), Lex30, WAT, VST

b. Criterion Variable: Narrative Total Score

c. Criterion Variable: Descriptive Total Score

d. Criterion Variable: Argumentative Total Score

Table 5 also reports that the results of ANOVA for the first model, ($F(3, 66) = 19.77, p = 0.00$) are considered as significant. The results are also significant for both second ($F(3, 66) = 18.07, p = 0.00$) and third ($F(3, 66) = 16.76, p = 0.00$) models. In other word, all three models including the combination of scores on VST, WAT and Lex30 can significantly predict EFL learners' total writing performances (narrative, descriptive, and argumentative task types respectively).

Table 6 indicates that the largest absolute β coefficient in the first, second and third model belongs to the scores on VST. WAT, on the other hand, is the second significant contributor of total narrative, descriptive and argumentative writing score. The examination of the Sig. values also indicates that Lex30 scores fail to significantly predict the criterion variables. However, the scores on VST and WAT make statistically significant unique contributions to the equation since their Sig. values are below .05.

Table 5

Regression Output: ANOVA for Models 1 to 3

Model ^d		Sum of Squares	df	Mean Square	F	Sig.
1 ^a	Regression	2624.73	3	874.91	19.77	.00
	Residual	2919.85	66	44.24		
	Total	5544.58	69			
2 ^b	Regression	2961.16	3	987.05	18.07	.00
	Residual	3604.20	66	54.60		
	Total	6565.37	69			
3 ^c	Regression	2287.82	3	762.60	16.76	.00
	Residual	3001.74	66	45.48		
	Total	5289.57	69			

a. Criterion Variable: Narrative Total Score

b. Criterion Variable: Descriptive Total Score

c. Criterion Variable: Argumentative Total Score

d. Predictors: (constant), VST, WAT, Lex30

Table 6

Regression Output: Coefficients for Model 1 to 3

Model		Unstandardized Coefficients			t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.46	6.33		1.65	.10
	VST	.28	.09	.38	3.13	.00
	WAT	.21	.06	.37	3.47	.00
	Lex30	.02	.06	.03	.30	.76
2	(Constant)	7.94	7.04		1.12	.26
	VST	.37	.10	.46	3.69	.00
	WAT	.19	.06	.31	2.82	.00
	Lex30	-.02	.07	-.03	-.28	.78
3	(Constant)	10.44	6.42		1.62	.10
	VST	.28	.09	.38	3.01	.00
	WAT	.18	.06	.33	2.98	.00
	Lex30	.02	.06	.04	.37	.70

1. Criterion Variable: Narrative Total Score

2. Criterion Variable: Descriptive Total Score

3. Criterion Variable: Argumentative Total Score

4.1.2. The Contribution of the Scores on VST, WAT and Lex30 to the Vocabulary Component Scores of Narrative, Descriptive, and Argumentative Writing

With regard to the second question, three standard multiple linear regressions were run to investigate whether the scores on VST, WAT and Lex30 predict the vocabulary component score of narrative, descriptive, and argumentative writing.

In Table 7, R Square for the fourth, fifth and sixth models are .63, .57 and .60 respectively specifying that these models, including the combination of the scores on VST, WAT and Lex30, can explain 63.40, 57.9 and 60 percent of the variance in vocabulary component score of the narrative, descriptive and argumentative respectively.

Table 7
Regression Model Summary- R and R Square

Model ^a	R	R Square	Adjusted R Square	Std. Error of the Estimate
4 ^b	.79	.63	.61	1.62
5 ^c	.76	.57	.56	1.77
6 ^d	.77	.60	.58	1.51

a. Predictors: (Constant), Lex30, WAT, VST

b. Criterion Variable: Narrative Vocabulary Score

c. Criterion Variable: Descriptive Vocabulary Score

d. Criterion Variable: Argumentative Vocabulary Score

Table 8 indicates that the results of ANOVA for the fourth $F(3, 66) = 38.06$, $p = 0.00$, fifth $F(3, 66) = 30.21$, $p = 0.00$ and sixth $F(3, 66) = 33.05$, $p = 0.00$ models are considered significant. This means that the combination of the scores on VST, WAT and Lex30 can significantly predict EFL learners' vocabulary component scores in narrative, descriptive, and argumentative writing task types

Table 8
Regression Output: ANOVA for Model 4 to 6

Model ^d		Sum of Squares	df	Mean Square	F	Sig.
4 ^a	Regression	301.11	3	100.37	38.06	.00
	Residual	174.03	66	2.63		
	Total	475.14	69			
5 ^b	Regression	286.27	3	95.42	30.21	.00
	Residual	208.44	66	3.15		
	Total	494.71	69			
6 ^c	Regression	228.98	3	76.32	33.05	.00
	Residual	152.38	66	2.30		
	Total	381.36	69			

a. Criterion Variable: Narrative Vocabulary Score

b. Criterion Variable: Descriptive Vocabulary Score

c. Criterion Variable: Argumentative Vocabulary Score

d. Predictors: (Constant), VST, WAT, Lex30

Table 9 also illustrates the Standardized Beta Coefficients in which VST and WAT in three models indicate statistically significant unique contributions to the equation because their Sig. values are below .05. However, Lex30 scores fail to show any significant contribution to the vocabulary component of narrative, descriptive and argumentative writing.

Table 9

Regression Output: Coefficients for Models 4 to 6

Models		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
4	(Constant)	-2.44	1.54		-1.58	.11
	VST	.12	.02	.55	5.45	.00
	WAT	.05	.01	.33	3.72	.00
	Lex30	.00	.01	.00	.03	.97
5	(Constant)	-1.46	1.69		-.86	.38
	VST	.12	.02	.56	5.16	.00
	WAT	.05	.01	.30	3.12	.00
	Lex30	-.00	.01	-.03	-.32	.74
6	(Constant)	-1.70	1.44		-1.17	.24
	VST	.09	.02	.48	4.53	.00
	WAT	.06	.01	.42	4.52	.00
	Lex30	-.01	.01	-.06	-.70	.48

4. Criterion Variable: Narrative Vocabulary Score

5. Criterion Variable: Descriptive Vocabulary Score

6. Criterion Variable: Argumentative Vocabulary Score

In table 9, the comparison of β values for the fourth, fifth and sixth models reveal that VST scores have the largest absolute β coefficient. The scores of WAT are the second significant predictor of vocabulary component score in narrative writing ($\beta = 0.33$, $t = 3.72$, $p = 0.00$), descriptive ($\beta = 0.30$, $t = 3.12$, $p = 0.00$) and argumentative ($\beta = 0.42$, $t = 4.52$, $p = 0.00$).

4.2. Discussion

This study explored the contribution of EFL learners' receptive knowledge of vocabulary breadth and depth and productive knowledge of vocabulary breadth measured by the VST, WAT and, Lex30 respectively, to their overall writing scores, and their vocabulary component scores in writing performance in three different writing tasks (descriptive, argumentative and narrative).

Regarding the first question, the findings of the study demonstrate that there was a moderately significant relationship between learners' breadth and depth of the receptive vocabulary knowledge with their narrative, descriptive and argumentative writing scores. This implies that learners

possessing extended knowledge of word breadth and depth are more skillful in their overall writing performance than those with smaller vocabulary size and depth. However, the results revealed that learners' breadth of productive vocabulary knowledge is less moderately but significantly correlated with their overall narrative, descriptive and argumentative writing scores. Our findings indicated that vocabulary knowledge, comprising receptive word knowledge, productive word knowledge and receptive vocabulary depth, can, altogether, significantly predict EFL learners' performances in overall narrative, descriptive, and argumentative writing, respectively. Nevertheless, according to the findings of this study, the learners' scores on VST are regarded the first predictor of their total descriptive and argumentative writing while the WAT scores are considered the second predictor. Moreover, Lex30 scores do not significantly predict the total scores of narrative, descriptive and argumentative writing.

Earlier studies have shown that the breadth of receptive vocabulary knowledge is, ranging from moderately to strongly, correlated with foreign language proficiency when evaluating writing, speaking, reading, and listening (Azodi et al., 2014; Johnson et al., 2016; Milton, 2013). Our results are consistent with Engber's (1995), Llach and Gallego's (2009) and Albrechtsen et al.'s (2008) studies. They found that learners' lexical proficiency is moderately correlated with their writing quality and learners with higher receptive word knowledge and lexical diversity achieved higher scores on their overall writings. Our results also tie in well with those in Schoonen et al.'s (2011) work. They pointed out that receptive vocabulary size is moderately correlated with descriptive writing performance. Our findings are also consistent with Stæhr's (2008) which indicated that writing ability is closely correlated with vocabulary size. Nation (2001) also argued that "vocabulary plays a significant role in the assessment of the quality of written work" (p.178). Schmitt (2014) stated that large vocabulary size is a special requirement for overall writing performance. The findings of our study enrich Milton's (2013) on the role of vocabulary size in writing performance. His results indicated that learners with higher proficiency level used a higher variety of words in their compositions than those with lower proficiencies.

This study supports Miralpeix and Muñoz's (2018) study which found a moderately significant relationship between X_Lex and Y_Lex, as the measures of receptive vocabulary size and writing performance scored by the holistic rubric. Our findings are also consistent with Dabbagh and Janebi's (2017) study which measured the breadth of receptive word knowledge by VLT, even though they found that the breadth of receptive word knowledge predicted descriptive writing performance assessed through Polio and Lim's (2020) revised analytic rubric.

Our findings are in line with Li and Kirby's (2014) study that found a significant and moderate relationship between EFL learners' summary writing performance and their depth of word knowledge. Dabbagh and Janebi (2017), partially in line with our study, concluded that learners' vocabulary depth is moderately but non-significantly correlated with overall descriptive writing performance. Albrechtsen et al. (2008) also operationalized the depth of word knowledge as network knowledge measured by WAT (with adjective samples). They investigated its relationship with overall writing ability based on TOEFL scoring guide, and our findings partially provide evidence for the results obtained by Albrechtsen et al. (2008) in that advanced EFL university students' scores on WAT are moderately but non-significantly correlated with their writing tasks. They provided some reasons for this non-significant relationship. Since there are subskills such as ideas, rhetorical features, arguments, and language control categories of skills which represent the writing product quality, the relationship between vocabulary depth and individual scores on these sub-skills is unlikely to be significant. Moreover, they suggested a need to incorporate a depth of vocabulary knowledge test with a wide variety of word classes and less frequent vocabulary in their study to reach a significant relationship between the depth of word knowledge and writing performance (Albrechtsen et al., 2008).

As to the second question, our results revealed a significant and moderate correlation between EFL learners' scores on WAT and VST as predictor variables and their scores on the vocabulary component in argumentative, narrative and descriptive writing tasks. However, scores on the Lex30 were less moderately but significantly correlated with the vocabulary component in narrative, descriptive and argumentative writing tasks. The scores of VST and WAT were, respectively, the first and second predictors of EFL learners' scores on vocabulary component in narrative, descriptive and argumentative writing tasks. Moreover, learner' Lex30 scores did not significantly predict the vocabulary component of narrative, descriptive and argumentative writing.

In this regard, the findings are consistent with earlier investigations which indicated that EFL learners' receptive vocabulary is the primary concern for them since lexical problems, together with poor word choice and less lexical diversity, derive from their small receptive vocabulary size (Engber, 1995; Llach & Gallego, 2009). Since VST is based on word frequency, it may predict EFL learners' lexical sophistication (i.e. using low-frequency words) as the main descriptor for the vocabulary component of writing performance of the three task types in the analytic rubric scale. Read (2000) stated that sophistication and lexical diversity contribute to L2 writing quality. Therefore, other results were broadly in line with Jeong (2017) and Fritz and Ruegg (2013), in that raters should be sensitive to the accuracy,

vocabulary size, lexical diversity as well as sophistication, when assessing essays. This study differed to some extent from Laufer's (1998) findings. She found a non-significant correlation between high-school learners' receptive vocabulary measured by VLT and lexical sophistication measured by the lexical frequency profile. However, lexical sophistication was assessed subjectively by two human raters in our study. Multiple test formats used to measure receptive vocabulary size may produce differences in the results. Apart from assessment differences, the current study included EFL undergraduate students studying non-English majors with individual learning differences which may be the cause of variation in results between Laufer (1998) and this study. This study's findings are also consistent with Webb's (2008) study which revealed that learners with more breadth of receptive word knowledge tend to use those words more than learners with a smaller mental lexicon.

Concerning the relationship between vocabulary depth and lexical sophistication, Bulté et al. (2008) found that lexical sophistication should be regarded as an indicator of qualitative word knowledge; they confirmed the positive correlation of vocabulary depth with the lexical richness of their investigated samples assessed by human raters. Depth of word knowledge may ease the use of more advanced words. Since vocabulary depth involves the networks between words in the mental lexicon, condensed networks may facilitate the retrieval of words. The students with poor vocabulary depth may retrieve the low-frequency words more slowly and decide on using accessible higher-frequency words. The fact that vocabulary component descriptors in the analytic scale imply that vocabulary diversity (use of non-repetitive words) is a secondary concern for scoring the vocabulary component encourages raters to assess the paradigmatic relations of vocabulary knowledge which is one aspect of measurement in WAT. Our results do not appear to corroborate Silverman et al. (2015) which did not reveal a significant correlation between the receptive aspect of vocabulary depth and word choice in a narrative writing task.

Our findings are not in agreement with Kilic's (2019) and Johnson, et al's (2016) study. They reported the contribution of the breadth of productive vocabulary knowledge in L2 writing performance. However, they used PVL T to measure controlled productive vocabulary as an aspect of productive vocabulary knowledge. Apart from test formats, learner proficiency levels and individual differences may change these findings.

5. Conclusion and Implications

This study aimed at determining the predictability of the breadth and depth of receptive vocabulary knowledge and breadth of productive vocabulary knowledge in the overall score and the vocabulary component of

EFL learners' narrative, descriptive, and argumentative writing performance. It concluded that the breadth and depth of receptive vocabulary knowledge significantly predicted both the overall writing and vocabulary component of the narrative, descriptive and argumentative writing tasks. It also concluded that the breadth of productive vocabulary correlated with the vocabulary component score as well as the total score of narrative, descriptive, and argumentative writing.

Raising learners' awareness about the breadth and depth of receptive vocabulary knowledge and their contributory roles in writing performance can be an important task for instructors. The learners' breadth of word knowledge should be enhanced in EFL classrooms by using meaningful activities to boost their vocabulary size for producing different writing task types. The learners' knowledge of semantics and word associations should also be developed either by syllabus materials and classroom teaching methodologies. It was found that students with more developed semantic networks are better at using vocabulary in their writing performance.

There are certainly some limitations to this investigation which limit its generalizability. This study needs to be replicated with a large number of participants and raters. Moreover, it provided 5-minute planning time for the learners before writing each task which may have influenced the rate of using words in each writing. The institute policy made us measure vocabulary tests in two sessions within one week. This time interval may have slightly affected learners' vocabulary size, although increases would have been too marginal, if at all. Another suggestion is also to replicate this research using the Vocabulary Knowledge Scale developed by Paribakht and Wesche (1997) as a measure for productive vocabulary depth construct. Only three genres of writing (descriptive, argumentative, and narrative) were considered in this research. It is advisable to investigate the correlation between other rhetorical modes of writing, such as critical, deductive, reflective, etc. to compare and contrast vocabulary knowledge along these dimensions. Furthermore, this study used one general topic for writing purpose. Other studies can use different topics for each rhetorical mode. Studying the correlation between learner's scores on Lex30 and their language learning skills like reading and listening can also be suggested for further research. Finally, researchers can provide writing tasks without planning time to the learners in order to prevent the influence of planning time on learners' use of more varied vocabulary items in their writing.

References

- Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System*, 38(3), 391-401.
- Alavi, S. M., & Akbarian, I. (2012). The role of vocabulary size in predicting performance on TOEFL reading item types. *System*, 40(3), 376-385.
- Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and Second language: Processes and development*. Palgrave Macmillan.
- Azodi, N., Karimi, F., & Vaezi, R. (2014). Measuring the lexical richness of productive vocabulary in Iranian EFL university students' writing performance. *Theory and Practice in Language Studies*, 4(9), 1837-1849.
- Bauer, L., Nation, P. (1993). Word families. *International Journal of Lexicography*, 6(4), 253-279.
- Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, 27(1), 101-118.
- 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.
- 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.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26(1), 42-65.
- Cheng, J., & Matthews, J. (2018). The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing*, 35(1), 3-25.
- Dabbagh, A., & Janebi Enayat, M. (2017). The role of vocabulary breadth and depth in predicting second language descriptive writing performance. *The Language Learning Journal*, 47(5), 8-11.
- Elgort, I. (2013). Effects of L1 definitions and cognate status on the vocabulary size test. *Language Testing*, 30(2), 253-272.
- Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journals of Second Language Writing*, 4, 213-238.
- Fitzpatrick, T., & Meara, P. (2004). Exploring the validity of a test of productive vocabulary. *Vigo International Journal of Applied Linguistics*, 1, 55-74.
- Fritz, E., & Ruegg, R. (2013). Rater sensitivity to lexical accuracy, sophistication and range when assessing writing. *Assessing Writing*, 18(2), 173-181.

- Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115, 523–547.
- 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. (2013). What vocabulary should we teach? Lexical Frequency profiles and lexical diversity in second language writing. *Writing and Pedagogy*, 5(1), 83–103.
- Johnson, M. D., Acevedo, A., & Mercado, L. (2016). Vocabulary knowledge and vocabulary use in second language 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.
- Kaivanpanah, S. & Zandi, H. (2009). The role of depth of vocabulary knowledge in reading comprehension in EFL contexts. *Journal of Applied Sciences*, 9(4), 698–706.
- Karami, H. (2013). The development and validation of a bilingual version of the vocabulary size test. *RELC Journal*, 43(1), 53–67.
- Kilic, M. (2019). Vocabulary knowledge as a predictor of performance in writing and speaking: A case of Turkish EFL learners. *PASAA: Journal of Language Teaching and Learning in Thailand*, 57(1), 1–33.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322.
- Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31(4), 537–561.
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language Learners over eight years of study: Single words and collocations. *EUROSLA Monographs Series*, 2(1), 127–148.
- Li, M., & Kirby, J. R. (2014). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, 36(5), 611–634.
- Llach, M. P. A., & Gallego, M. T. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Journal of the Spanish Association of Anglo-American Studies*, 31(1), 129–147.
- Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, 72, 23–36.
- Meara, P., & Fitzpatrick, T. (2000). Lex 30: An improved method of assessing productive vocabulary in an L2. *System*, 28, 19-30.
- Meara, P., & Jones, G. (1990). *Eurocentres Vocabulary Size Test, Version E1.1/K10*. Zurich: Eurocentres Learning Service.
- Meara, P., & Wolter, B. (2004). V-Links: Beyond vocabulary depth. In Albrechtsen, D., Haastруп, K., & Henriksen, B. (Eds.), *Writing and*

- vocabulary in foreign language acquisition* (pp. 85–96). Museum Tusculanum Press.
- Mehrpour, S., Razimjo, S.A. & Kian, P. (2011). The relationship between breadth and depth of vocabulary knowledge and reading comprehension among Iranian EFL learners. *Journal of English Language Teaching and Learning*, 2(222), 97-127.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer. (Eds.), *L2 vocabulary acquisition, knowledge, and use: New perspectives on assessment and corpus analysis*, (pp. 57–78). Eurosla Monographs Series, Rome, Italy: Eurosla.
- Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in English as a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, M Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 83-98). Toronto: Multilingual Matters.
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), 1–24.
- Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical frequency profiles across drafts. *System*, 30(2), 225–235.
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing use and success strategy. *The Modern Language Journal*, 90(3), 387–401.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nguyen, L. T. C., & Nation, I. S. P. (2011). A bilingual vocabulary size test of English for Vietnamese learners. *RELC Journal*, 42(1), 86–99.
- Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing: An Interdisciplinary Journal*, 26(1), 45–65.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In Coady, J., & Huckin, T. (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200), Cambridge University Press.
- Polio, C. & Lim, J. (2020). Revising a writing rubric based on raters' comments: Does it result in more valid and reliable assessment? In M. Riazi, L. Shi, & K. Barkaoui (Eds.), *Writing in a second language: Learning, teaching, and assessment: Studies in honor of Alister Cumming*. Cambridge Scholars Press.

- Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish– English bilingual children in elementary school. *Reading and Writing, 25*(7), 1635–1664.
- Qian, D. D. (1998). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*(2), 282–307.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10*(3), 355–371.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies, 7*(2), 105–125.
- Richards, J.C. (2008) *Teaching listening and speaking*. 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*(1). <https://doi.org/10.1186/2229-0443-3-12>
- 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*(4), 1–19.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951.
- Schoonen, R., van Gelderen, A., Stoel, R. D., Hulstijn, J., & de Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning, 61*(1), 31–79.
- Silverman, R. D., D. Coker, C. P. Proctor, J. Harring, K. W. Piantedosi, & A. M. Hartranft. (2015). The relationship between language skills and writing outcomes for linguistically diverse students in upper elementary school. *The Elementary School Journal, 116*(1), 103–125.
- Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal, 36*(2), 139–152.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.
- Teng, F. (2016). An in-depth investigation into the relationship between vocabulary knowledge and academic listening comprehension. *TESL-EJ, 20*(2), 1–17.
- Uchihara, T., & Saito, K. (2019). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal, 47*(1), 64–75.

- Walters, J. (2012). Aspects of validity of a test of productive vocabulary: Lex30. *Language Assessment Quarterly*, 9(2), 172–185.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30(1), 79–95.
- 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.
- Yoon, H., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275–301.