The Interaction Between Lexical Cohesion and L2 Vocabulary Size and Depth in TOEFL Reading Item Types

Mostafa Janebi Enayat[1] iD *& Farhad Mazlum[2]* iD

Abstract

Few studies have addressed the interaction among reader and text variables in reading comprehension. This research probed the role of aspects of lexical knowledge to the TOEFL reading comprehension of passages with low and high lexical cohesion. The extent of this contribution was checked for five types of TOEFL reading test items: Main Idea, Stated Detail, Inference, Reference, and Lexical Inferencing. For this purpose, 60 Iranian English as a Foreign Language (EFL) took the New Vocabulary Levels Test (NVLT), Word Associates Test (WAT), and TOEFL reading comprehension tests with low and high lexical cohesion determined through the computational tool Coh-Metrix. Results of Wilcoxon signed-rank tests revealed that the students outperformed on three types of TOEFL reading items (Main Idea, Inference, and Lexical Inferencing) for texts with high lexical cohesion. Multiple linear regression analyses further showed that vocabulary depth was a stronger predictor of the participants' scores on Lexical Inferencing item of TOEFL reading texts with high lexical cohesion, while vocabulary size predicted the same item performance for the low-lexical-cohesion subtest. Moreover, vocabulary depth could significantly predict performance on Main Idea test item of TOEFL high-lexical-cohesion subtest over and above the low-lexical-cohesion subtest. The implications for teachers, test designers, and materials writers are discussed.

Keywords: vocabulary size, vocabulary depth, TOEFL reading, reading test item types

[1] Corresponding Author: English Department, Faculty of Humanities, University of Maragheh, Maragheh, Iran; *Email: m.enayat@maragheh.ac.ir*; ORCID ID: http://orcid.org/0000-0001-8724-9121
[2] English Department, Faculty of Humanities, University of Maragheh, Maragheh, Iran; *Email: mazlumzf@yahoo.com*; ORCID ID: http://orcid.org/0000-0002-4955-4472

1. Introduction

The factors capable of affecting performance in language tests can be classified into three main sets: (1) task-related features; (2) attributes that are related to the test-takers; and (3) associations between characteristics of task and learners (Bachman, 2002). Alderson (2000) highlighted that there exist various factors related to the text or task that can either aid or impede the process of comprehending a text. These factors encompass aspects such as the content, types or genres, organization, structure, vocabulary, typography, layout, the interaction between verbal and nonverbal text, and the medium through which the text is delivered. Among test takers' characteristics that affect test performance are L1, prior knowledge, affective and cognitive attributes, ethnic background, and gender (Bachman & Palmer, 1996).

Numerous investigations have made attempts to examine the impact of textual characteristics and reader factors on the proficiency of individuals taking text-dependent assessments, including reading evaluations, cloze tests, and C-tests. They ranged from examining the text difficulty (e.g., Babaii & Jalali Moghaddam, 2006; McDaniel et al., 2002), text organization (e.g., Frestl & Von Cramon, 2001; MacMillan, 2007; McNamara, 2001; Ozuru et al., 2009), to reader variables like prior knowledge (e.g., Calisir & Gurel, 2003; Nguyen et al., 2022; Potelle & Rouet, 2003), reading skill (e.g., Alavi & Akbarian, 2012; Jackson, 2005; Kaivanpanah et al., 2022; Ozuru et al., 2009), grammatical knowledge (e.g., Shiotsu & Weir, 2007), and vocabulary knowledge (e.g., Janebi Enayat & Derakhshan, 2021; Matthews, 2018; Nassaji, 2006; Qian, 2002; Zhang & Anual, 2008). However, the interaction between reader and text variables (Bachman, 2002) is rarely addressed in the literature despite the fact that the former could contribute to the latter (Babaii & Jalali Moghaddam, 2006; Janebi Enayat & Babaii, 2018).

One of the key reader variables that is evidenced to affect context-based tests is vocabulary knowledge (Read & Chapelle, 2001) since contextual clues like lexical chains available in text-dependent tests (e.g., C-test, cloze test, and reading) could be realized by more lexically proficient test-takers (Janebi Enayat & Babaii, 2018; Nassaji, 2006). The dimensions of vocabulary size (i.e., number of L2 words that learners know) and vocabulary depth (i.e., quality of vocabulary knowledge that

includes knowing word associations and collocation) are among the lexical variables that could correlate with lexical features of text like lexical cohesion as both are linked with word associations (Qian, 2004). Additionally, as explicated by Meara and Wolter (2004), more vocabulary size would enhance vocabulary depth as learning more words could develop the lexical network. Taking into account the essential role of vocabulary knowledge in receptive skills and the importance of text variables in language instruction and learning, this study aimed at probing the role of size and depth aspects of vocabulary knowledge in TOEFL reading comprehension of texts with high and low lexical cohesion. The significance of vocabulary in language proficiency testing, particularly in the context of reading comprehension, is firmly established (Alavi & Akbarian, 2012). For instance, in the development of the TOEFL Internet-based Test (iBT), vocabulary is considered a crucial element (Qian, 2008). The significance of this factor is also validated in the TOEFL 2000 Reading Framework (Enright et al., 2000). Consequently, the reading section of TOEFL comprises three passages, each consisting of 650 to 750 words, with all passages incorporating vocabulary items. The role of lexical cohesion and aspects of vocabulary knowledge in the performance of five reading test items, namely main idea, stated detail, inference, reference, and lexical inferencing, was further probed. As delineated in prior research (Alavi & Akbarian, 2012; Nassaji, 2006), distinct reading assessment items necessitate varied dimensions of lexical competence, yet the correlation between these item types and lexical cohesion remains unexplored in existing literature.

Vocabulary and reading are associated with each other as Nassaji (2006) argued that attention to different aspects of lexical knowledge can assist the employment of lexical features of reading comprehension passages. Previous studies have found significant associations between knowledge of vocabulary and reading performance (e.g., Alavi & Akbarian, 2012; Janebi Enayat & Babaii, 2018; Nassaji, 2006; Qian, 2002; Stæhr, 2008; Tong & Tong, 2022). However, these studies merely focused on reader variables, failing to delineate the interaction between reader and text variables. Moreover, the contribution of lexical cohesion to the test-takers' reading performance has not been properly addressed in L2 research. The present study, thus, endeavors to investigate the function of lexical connections, as textual variables, and their

association with dimensions of vocabulary knowledge, as factors associated with readers, in TOEFL reading items of texts with both high and low semantic associations. This investigation examines both the size and depth of vocabulary knowledge since, to genuinely grasp the lexical proficiency of an L2 learner, we must surpass basic examinations of lexical elements and commence the analysis of the language learner's understanding of semantic relations among words (Crossley & McNamara, 2009). Given the importance of reader variables like vocabulary size and depth, and lexical cohesion as a text variable, in TOEFL reading, two research questions were addressed:

1. Does lexical cohesion make a difference in five types of TOEFL reading comprehension test items?

2. Which aspect of vocabulary knowledge could significantly contribute to the use of lexical cohesion to answer different test item types of TOEFL reading?

2. Literature Review

2.1. Vocabulary Knowledge

Vocabulary knowledge is not a binary term, but rather encompasses varying degrees of comprehension (Shen, 2008). There are two primary facets to vocabulary knowledge: the receptive and productive aspects. According to Mondria and Wiersma (2004), receptive vocabulary knowledge entails understanding the meaning of a word in a second language and being able to translate it into the first language. On the other hand, productive vocabulary knowledge involves the ability of language learners to translate a word from their first language into the second language. Henriksen (1999) further delineated three dimensions of vocabulary knowledge: (a) the range of knowledge from partial to precise, (b) the depth of knowledge, and (c) the ability to transition from receptive to productive use. The first dimension is concerned with accurate comprehension, which aids learners in translating a word into their first language, selecting the correct option in a multiple-choice test item, or paraphrasing in the target language. The second aspect, vocabulary depth, encompasses the quality of lexical knowledge, including word associations and collocations. The final dimension

pertains to the receptive and productive aspects, which refer to the understanding of a word and the ability to use it in written or spoken language.

Vocabulary size and depth have been identified as predictor variables of both the receptive and productive skills (e.g., Alharthi, 2020; Dabbagh & Janebi Enayat, 2019; Derakhshan & Janebi Enayat, 2020; Janebi Enayat & Derakhshan, 2021; Koizumi & In'nami, 2013; Matthews, 2018; Matthews & Cheng, 2015; Uchihara & Clenton, 2020; Uchihara & Saito, 2019). The semantic associations that these variables entail could provide a valid justification for their effects (Janebi Enayat & Babaii, 2018) – while vocabulary size refers to the syntagmatic lexical relations that could co-occur in a sentence (e.g., room-bed and room-dark), paradigmatic relations are hierarchical (e.g., sparrow-bird and bird-wing) (Schoonen & Verhallen, 2008). The use of such lexical relations in writing and speaking, for example, could result in a more lexically infrequent and cohesive language (Dabbagh & Janebi Enayat, 2019; Janebi Enayat & Derakhshan, 2021; Koizumi & In'nami, 2013; Uchihara & Clenton, 2020). In other words, such lexical variables are associated with text features that could ultimately determine the test-takers' performance.

2.2. Lexical Cohesion

When reading a written composition, whether it be for amusement or for the purpose of completing a task such as filling in the missing words in a C-test or answering comprehension questions, we do not perceive the text as a mere sequence of words. Instead, subconsciously, we endeavor to establish connections between these words in order to achieve a comprehensive understanding of the content (Alshaar, 2008). One aspect of text that a learner can utilize to aid their comprehension is cohesion and coherence. In a seminal study, Halliday and Hasan (1976) classified the text feature of cohesion into two kinds: grammatical and lexical. The former encompasses elements like substitution, reference, conjunction, and ellipsis, while the latter is subdivided into collocation and reiteration. Collocation involves the co-occurrence of lexical units, whereas reiteration pertains to the repetition of a word, either verbatim or as a synonym.

Hoey (1991) proposed a system to characterize lexical patterns that was grounded

in the idea of link. The term *link* was chosen over the traditional term *tie* employed by Halliday and Hasan (1976) because *tie* implies directionality whereas *links* implies multidirectionality, thus permitting the formation of webs among lexical items (Sardinha, 2001). Halliday and Hasan (1976) introduced the concept of a chain to depict a relationship where one element refers to another, and so on. Nevertheless, Morris and Hirst (1991) argued that the concept of lexical cohesion extends beyond word pairs to encompass a series of closely related words within a textual unit. They labeled these chains of related words as lexical chains, formed based on semantic connections between words. Silber and McCoy (2002) highlighted that these semantic relations are united by semantically linked words like superordinates/subordinates, synonyms, and hypernyms/hyponyms.

2.3. Relationship between Lexical Cohesion and Vocabulary Size and Depth

The establishment of lexical chains or connections stems from the semantic associations present among words, as previously mentioned. These associations consist of synonyms, hypernyms/hyponyms, as well as superordinates/subordinates, according to Silber and McCoy (2002). The concept is in line with the meaning components of depth of lexical knowledge, which pertains to an understanding of the syntagmatic and paradigmatic relationships among words, as noted by Read (1993). Syntagmatic and paradigmatic relations are established through semantic links like synonymy, hyponymy, meronymy, and collocation. Studies have shown that the utilization of these associations aids in the process of lexical inferencing, as seen in Janebi Enayat and Babaii (2018), Nassaji (2006), and Qian (2004). Thus, this research suggests that the depth of vocabulary knowledge, as a factor related to the reader, can impact the test takers' utilization of lexical cohesion during text comprehension. This assumption is based on the shared involvement of syntagmatic and paradigmatic relations in both variables. To summarize, learners' ability to partake in lexical inferencing and employ contextual cues may be influenced by their depth of vocabulary knowledge.

Additionally, the size and depth aspects of vocabulary knowledge are associated

(Dabbagh & Janebi Enayat, 2022; Schmitt, 2014) as the former develops the latter through building more connections between and among lexical items (Meara & Wolter, 2004). In other words, learning more words (vocabulary size) would create a stronger lexical network (vocabulary depth) in which the new lexical items could be semantically associated. This interconnection between size and depth of vocabulary is well-established in literature as many scholars have found significant correlations between scores of these dimensions of vocabulary knowledge (e.g., Akbarian, 2010; Dabbagh & Janebi Enayat, 2022; Henriksen, 2008; Janebi Enayat & Amirian, 2020; Janebi Enayat et al., 2018; Schmitt, 2014). Accordingly, without a sufficient vocabulary size, contextual clues are rendered ineffective (Janebi Enayat & Babaii, 2018; Kaivanpanah & Alavi, 2008) since the known words would assist the test-takers guess the unknown ones, and if a large portion of a reading passage, for instance, is lexically difficult, lexical inferencing would seem pointless (Nassaji, 2006; Nation, 2011). One of the contextual clues that could be employed by the text-dependent test-takers are lexical chains or lexical cohesion. This study, therefore, hypothesized that vocabulary size and depth, which are interrelated aspects of vocabulary knowledge, could be influential in assisting the students to benefit from lexical cohesion of the passages used in reading tests.

2.4. Review of Related Studies

Vocabulary size and depth have been recognized as key determiners of L2 learners' success in reading comprehension. Considering cloze passages as measures of reading comprehension, Schoonen and Verhallen (1998) examined the role of size and depth of vocabulary knowledge in reading comprehension ability of Dutch primary school children. The results indicated that both aspects of vocabulary knowledge had a 5-10% contribution to the prediction of cloze passages. This study used cloze passages, which rely on dependent test items, to examine the role of only reader variables while the investigation of text features like lexical and syntactic difficulty of the cloze passages could provide a more insightful view of the interaction between aspects of vocabulary knowledge and cloze test performance. A similar study using conventional reading comprehension tests that employ independent items

was conducted by Shiotsu and Weir (2007) who delved into investigating the interplay among syntactic knowledge, vocabulary breadth, and L2 reading performance, with a specific emphasis on the predictive capacity of these variables. By utilizing Structural Equation Modeling (SEM) as a methodological approach, the scholars underscored the importance of syntactic knowledge compared to vocabulary size in influencing outcomes on reading comprehension evaluations. Conversely, Zhang and Anual (2008) identified that vocabulary size played a pivotal role in forecasting reading comprehension achievements, especially at the 2000-word and 3000-word thresholds among student cohorts. A couple of studies, however, probed the extent that vocabulary knowledge would assist the test-takers in lexical inferencing. Qian (1998), for instance, investigated how size and depth of word knowledge were related to the 74 English as a foreign language (EFL) learners' comprehension of general academic texts. Furthermore, the predictive ability of these aspects in reading comprehension was examined. The results of his research revealed that (a) the two aspects of size and depth were positively correlated with reading comprehension scores, (b) vocabulary depth contributed to reading comprehension over and above size of vocabulary knowledge, and (c) students with more depth of lexical knowledge made use of lexical inferencing and focused more on word meanings rather than word forms. Following his previous research, Qian (2002) sought to ascertain the contribution of vocabulary size and certain factors pertaining to vocabulary depth (namely, synonymy, polysemy, and collocation) to fundamental reading comprehension. Furthermore, he aimed to assess the utility of a vocabulary size metric and an adapted vocabulary depth metric in the evaluation of reading comprehension. The outcomes of his investigation provided substantiation for the intimate interconnection between the scope and depth of vocabulary knowledge and reading performance. In summary, comparable to his prior research, components of vocabulary were discovered to be indicative of reading comprehension. Additionally, vocabulary depth was determined to access a more profound understanding of vocabulary compared to vocabulary size. Moreover, the findings indicated that employing a combination of vocabulary size and depth could more potently forecast reading performance. These studies were a step forward in identifying the role of vocabulary knowledge in answering certain reading comprehension items like lexical

inferencing, yet the text features were not addressed and only one type of test item (i.e., lexical inferencing) was examined.

Qian (1998) previously established that a deeper vocabulary knowledge leads to the application of lexical inferencing for enhancing reading comprehension. Nassaji (2004) and Nassaji (2006) continued this line of investigation, exploring the connection between the depth of vocabulary knowledge among English as a Second Language (ESL) learners, the usage of lexical inferencing strategies, and the capability to infer word meanings from context. Through the administration of a Word Associates Test (WAT) for assessing vocabulary depth, a reading comprehension test, and an introspective think-aloud technique, the scholar's research revealed that (a) students with greater vocabulary depth tended to employ lexical inferencing more than those with limited vocabulary depth, and (b) the depth of word knowledge significantly contributed to inferential success beyond the learners' strategy employment levels. The two investigations undertaken by Nassaji (2004, 2006) made significant contributions to the understanding of how vocabulary size and depth assist L2 learners comprehend reading passages, particularly the comprehension items that require the use of lexical clues for inferring the meaning of unknown words. However, the lexical features of passages and the extent that other comprehension items would interact with them were not addressed.

The only study that focused on the interaction between performance on TOEFL reading items (i.e., Main Idea, Stated Detail, Inference, Reference, and Lexical Inferencing) and vocabulary size was Alavi and Akbarian (2012) who found that vocabulary size was correlated with performance in three of the TOEFL test items (Main Idea, Lexical Inferencing, and Stated Detail) and only one test item (Lexical Inferencing) was correlated with vocabulary size for the highest proficiency students, suggesting that the inclusion of any type of reading test item regardless of the students' vocabulary knowledge or language proficiency would be ineffective. Likewise, this study failed to delineate the interaction between size and depth aspects of vocabulary knowledge and lexical features of text in the performance of different TOEFL reading test items.

In an attempt to explore the association between reader and text variables,

however, Janebi Enayat and Babaii (2018) delved into the intricate connection between vocabulary size and depth, and lexical cohesion in the context of reading comprehension. By utilizing C-test as a text-dependent instrument, the researchers explored how vocabulary size and depth influenced the utilization of lexical bonds in C-test processing among Iranian EFL learners. Drawing on Hoey's (1991) lexical analysis, the authors categorized C-tests into high- and low-bond texts to examine the impact of vocabulary depth on high-bond C-test performance, revealing a significant predictive relationship which indicated that the test-takers with higher vocabulary depth could benefit more from lexical clues in the C-test passages that contained higher lexical bonds.

As evident in the review of the related studies, the interaction between reader and text variables in general, and the extent that size and depth aspects of vocabulary knowledge could be associated with the lexical cohesion of text in the context of TOEFL reading test items has not been addressed in L2 research. The investigation of this interaction in a high-stakes test that is widely used around the globe would offer valuable insights to stakeholders regarding the specific aspects of vocabulary knowledge and test item types that warrant emphasis in TOEFL reading preparation courses.

3. Methodology

3.1. Participants

A total of 60 upper-intermediate, lower-advanced, and upper-advanced Iranian EFL students (18 males and 42 females) in the age range of 18-27 (mean age = 21.24) participated in this study. The participants' general English language proficiency had been homogenized using Oxford Quick Placement Test (OQPT). All the students were native speakers of Azeri who were selected from three private language institutes using convenience random sampling.

3.2. Instruments

Oxford Quick Placement Test evaluates the overall language proficiency of individuals. It includes 60 items focusing on vocabulary, grammar, and reading comprehension. Geranpayeh (2003) reported Cronbach's alpha reliability of 0.90 for this test.

Word Associates Test (WAT) was designed by Read (1993) to assess vocabulary depth. Its structure takes into consideration the primary associations of paradigmatic, syntagmatic, and analytic connections between two words. Since its development, the test has undergone revisions based on validation research (e.g., Ishii & Schmitt, 2009; Schmitt et al., 2011). The test has been reported to yield acceptable estimates of reliability (e.g., Janebi Enayat et al., 2018; Janebi Enayat & Derakhshan, 2021; Qian, 2002).

The New Vocabulary Levels Test (NVLT) was established by Webb et al. (2017) to assess vocabulary size across various word-frequency levels such as 1,000, 2,000, 3,000, 4,000, and 5,000. The initial lists primarily drew from spoken text rather than written text to suit students learning English as a foreign language (Webb & Sasao, 2013).

TOEFL Reading Subtests were selected for the purpose of this study: one using texts with high lexical cohesion and one with low lexical cohesion. The subtests were selected from official paper-based tests published by Educational Testing Service (ETS) (TOEFL ACTUAL TESTS, 2005) for the sake of reliability. For the purpose of this study, five reading comprehension test items were selected: (1) guessing the meaning of unknown words using contextual clues (henceforth, Lexical Inferencing), (2) getting the gist or main idea of a passage (henceforth, Main Idea), (3) locating the stated details or factual information (henceforth, Stated Detail), (4) finding references of pronouns (henceforth, Reference), and (5) inferring the unstated information like the author's tone or perspective (henceforth, Inference).

Coh-Metrix. The advanced computational tool of Coh-Metrix measures lexical overlap between sentences and paragraphs (McNamara et al., 2010; McNamara, Graesser et al., 2014). This tool automatically processes texts and reports indices of cohesion. Latent Semantic Analysis (LSA) is used by Coh-Metrix to measure the semantic relations between sentences or paragraphs (Landauer et al., 2007). Three

indices of LSA were used to measure the lexical cohesion of TOEFL reading passages: (1) LSA for adjacent sentences, (2) LSA for all sentences in paragraph, and (3) LSA for adjacent paragraphs. Each index varies from 0 (low cohesion) to 1 (high cohesion).

3.3. Data Collection and Analysis

To collect the data, the OQPT was first administered to 115 students who had 45 minutes to take this placement test. Those who obtained a score between 36-60 (n = 60) were selected for the main phase of this study. The students then took the NVLT and WAT in 60 minutes. The first three tests were administered in one seating in a controlled testing environment where invigilators were present and cheating was not allowed. Exact scoring method was used for these tests. As for TOEFL reading subtests with low and high lexical cohesion, the passages that seemed to have higher or lower lexical cohesion were typed and analyzed using the Coh-Metrix computational tool. A subtest with LSA indices below 0.55 was used as the TOEFL reading with low lexical cohesion and a subtest with LSA indices above 0.65 was selected as the subtest with high lexical cohesion (Crossley et al., 2016a). In the final test administration, which was held after a one-week interval, the two subtests of TOEFL reading were administered in 90 minutes.

SPSS version 27.0 was used to analyze the obtained data. To compare the participants' reading scores on the two sub-tests of TOEFL reading with low and high lexical cohesion (RQ1), the non-parametric Wilcoxon signed-rank test was run. Multiple regression analyses were employed to probe the role of size and depth aspects of vocabulary knowledge to the use of lexical cohesion in high- and low-cohesion reading tests.

4. Results

4.1. Normality of Data and Descriptive Statistics

Before running the inferential tests for the tests and subtests of the study, normality

of the data was checked and descriptive statistics of the data were obtained. Table 1 shows that the data are not normally distributed for the sets of scores (*p*<.05). Hence, the non-parametric Wilcoxon signed-rank tests were used for the first research question. Descriptive statistics show that the students' mean score on WAT was relatively higher than the NVLT. Moreover, the test takers' mean scores on the reading test items of the high-lexical-cohesion subtest (HLCS) were higher than the low-lexical-cohesion subtest (LLCS).

Table 1

The Results of Normality and Descriptive Statistics for the Scores of Vocabulary Size and Depth as well as the TOEFL Reading Items Across High- and Low-Lexical-Cohesion Subtests (n = 60)

	Shapiro-Wilk			Mean	*SD*	Range	*MPS*
	Statistic	*df*	Sig.				
NVLT	.937	60	.004	51.22	13.31	54.00	150
WAT	.959	60	.043	64.71	11.69	48.00	100
Main idea (HLCS)	.795	60	.000	3.13	.79	2.00	4
Main idea (LLCS)	.782	60	.000	1.93	.63	2.00	4
Stated detail (HLCS)	.901	60	.000	9.45	1.19	5.00	11
Stated detail (LLCS)	.880	60	.000	9.40	1.19	6.00	11
Inference (HLCS)	.427	60	.000	2.85	.36	1.00	3
Inference (LLCS)	.745	60	.000	1.58	.64	2.00	3
Reference (HLCS)	.850	60	.000	2.06	.86	3.00	4
Reference (LLCS)	.840	60	.000	2.05	.94	3.00	4
LI (HLCS)	.948	60	.012	13.53	2.81	10.00	19
LI (LLCS)	.925	60	.001	6.46	2.28	8.00	19

Note: HLCS = High-lexical-cohesion subtest; LLCS = Low-lexical-cohesion subtest; LI = Lexical inferencing; SD=Standard deviation; MPS = Maximum possible score.

4.2. The Contribution of Lexical Cohesion to TOEFL Reading

The first research question examined the effect of having high and low lexical cohesion in the texts used for TOEFL reading test items. The non-parametric Wilcoxon signed-rank test was used to find the differences between the scores obtained for the five item types of TOEFL reading. The results, as shown in Table 2, indicate that lexical cohesion could significantly affect the students' reading scores on three types of items: Main Idea ($Z = -5.646$, $p < .001$), Inference ($Z = -5.646$, $p < .001$), and Lexical Inferencing ($Z = -5.646$, $p < .001$). The results unveiled the highly

significant contribution of lexical cohesion to the reading performance of students on three TOEFL reading item types.

Table 2
Related-Samples Wilcoxon Signed-Rank Test Results for the Contribution of Lexical Cohesion to TOEFL Reading Item Types

Item type	N	TS	SE	Z	p	Decision
Main Idea	60	.000	97.40	-6.28	.001***	Reject the null hypothesis
Stated Detail	60	585.00	95.61	-.031	.975	Retain the null hypothesis
Inference	60	31.00	115.57	-6.39	.001***	Reject the null hypothesis
Reference	60	629.00	97.40	-.087	.930	Retain the null hypothesis
Lexical Inferencing	60	.000	135.24	-6.76	.001***	Reject the null hypothesis

Note: TS=Test Statistic; SE=Standard Error
***$p<.001$.

4.3. The Contribution of Vocabulary Size and Depth to the Use of Lexical Cohesion in TOEFL Reading Item Types

The second research question of this study investigated the contribution of size and depth of vocabulary knowledge to TOEFL reading item types of texts with high and low lexical cohesion. For this purpose, multiple linear regression analysis (using the stepwise method) was run. The collinearity statistics proved that multicollinearity was not observed among the predictor variables since the tolerance values were below 0.40 and the variance inflation factors were below 10 (see Table 3) (Tabachnick & Fidell, 2012). The assumptions of linearity and homoscedasticity were also checked. The results indicated that no variables were entered into the equation for the three item types of Stated Detail, Inference, and Reference. The results for the Main Idea reading item for the HLCS, as shown in Table 3, revealed that two models emerged for this relationship. The first model in which only the WAT was entered as the predictor variable could explain about 72% of the variance in the dependent variable ($F (1,58) = 151.506$, $p < .001$, $R^2 = .723$). The second model where both WAT and NVLT were entered as the explanatory variables could explain 75% of the Main Idea reading test item performance in the HLCS ($F (2,57) = 86.956$, $p < .001$, $R^2 = .753$). In other words, the addition of the NVLT scores could provide an additional 3% of the predictive power which was a significant change ($p < .05$).

Table 3

Multiple Regression Analyses for Vocabulary Size and Depth in TOEFL Reading of Texts with High Lexical Cohesion (Main Idea Test Item)

	R	R^2	ΔR^2	Unstandardized		Standardized	Collinearity statistics	
				B	SE B	β	Tolerance	VIF
Model 1	.850	.723***						
Constant				-.590	.307			
WAT				.058	.005	.850***	1.000	1.000
Model 2	.868	.753***	.030*					
Constant				-.909	.317			
WAT				.080	.010	1.186***	.210	4.756
NVLT				-.022	.009	-.378*	.210	4.756

*$p < 0.05$, ***$p < 0.001$.

Another multiple linear regression analysis (using the stepwise method) was run for the Main Idea test item of LLCS. The results provided in Table 4, revealed that just one model emerged for this relationship in which only the WAT was entered as the predictor variable. It could explain about 39% of the Main Idea reading test item performance in the LLCS (F (1,58) = 38.782, $p < .001$, R^2= .394).

Table 4

Multiple Regression Analyses for Vocabulary Size and Depth in TOEFL Reading of Texts with Low Lexical Cohesion (Main Idea Test Item)

	R	R^2	Unstandardized		Standardized
			B	SE B	β
Model 1	.628	.394***			
Constant			-.271	.364	
WAT			.034	.006	.628***

***$p < 0.001$.

The results also revealed that performance on the Lexical Inferencing test item of HLCS was significantly associated with depth of vocabulary knowledge (WAT scores) (Table 5). The only model in which the WAT was identified as the predictor variable could explain about 78% of the variance in the dependent variable (F (1,58) = 207.449, $p < .001$, R^2= .782).

Table 5

Multiple Regression Analyses for Vocabulary Size and Depth in TOEFL Reading of Texts with High Lexical Cohesion (Lexical Inferencing Test Item)

	R	R^2	Unstandardized		Standardized
			B	SE B	β
Model 1	.884	.782***			
Constant			-.228	.971	
WAT			.213	.015	.884***

***$p < 0.001$.

The results of multiple linear regression analysis (using the stepwise method) for the Lexical Inferencing test item of LLCS revealed that just one model emerged for this association in which the only predictor variable was vocabulary size (NVLT scores) (Table 6). It could explain about 75% of the Main Idea reading test item performance in the LLCS (F (1,58) = 178.805, $p < .001$, $R^2 = .755$).

Table 6

Multiple Regression Analyses for Vocabulary Size and Depth in TOEFL Reading of Texts with Low Lexical Cohesion (Lexical Inferencing Test Item)

	R	R^2	Unstandardized		Standardized
			B	SE B	β
Model 1	.869	.755***			
Constant			-1.185	.591	
NVLT			.149	.011	.869***

***$p < 0.001$.

5. Discussion

This study tried to identify the interaction between size and depth aspects of vocabulary knowledge and lexical cohesion in TOEFL reading test items. The effect of having high lexical cohesion in reading comprehension was probed for five test item types of TOEFL reading: Main Idea, Stated Detail, Inference, Reference, and Lexical Inferencing. Moreover, the contribution of vocabulary size and depth to the use of lexical cohesion in performance on these reading test items was investigated. The results of Wilcoxon signed-rank tests and multiple linear regression analyses for the scores of 60 EFL students indicated that the TOEFL reading comprehension of

passages with high lexical cohesion was significantly higher for three test items: Main Idea, Inference, and Lexical Inferencing. Additionally, vocabulary depth was a stronger predictor of the scores on Lexical Inferencing and Main Idea test items in HLCS, whereas vocabulary size was a stronger explanatory variable for the Lexical Inferencing test item in LLCS.

The results of the first research question showed a significant association between lexical cohesion and TOEFL reading comprehension. The findings showed that having higher lexical cohesion in texts lead to better reading comprehension, particularly for Main Idea, Inference, and Lexical Inferencing test items. This is in line with previous studies which found a positive effect for text comprehension when text variables like lexical cohesion are applied and considered (Janebi Enayat & Babaii, 2018). This finding could be justified with reference to the lexical clues that are provided by means of lexical cohesion in the text. Such clues are provided by semantic associations among words that contribute to the overall comprehension of the text by helping the students do lexical inferencing and make meaningful relationships among sentences. Sigott (2004) contends that lexical chains available in a context-dependent test could assist the test-takers in decoding the unknown text. The contextual clues like lexical links and semantic associations among sentences and paragraphs could improve the macro-level processing of text, which would lead in higher comprehension (Babaii & Fatahi-Majd, 2014). This study provided evidence for the positive effects of lexical cohesion in answering TOEFL reading test items like Main Idea, Inference, and Lexical Inferencing, which are corroborated by previous studies (e.g., Babaii & Ansary, 2001; Hastings, 2002; Janebi Enayat & Babaii, 2018; Nassaji, 2006). Additionally, the results are in line with Babaii and Jalali Moghaddam (2006) which found that lexical chains in context-based tests could be employed by the test-takers to fill in the mutilated words of a C-test. This signals the importance of contextual clues in reading exams where the test-takers are advised to look for lexical clues that could assist them in both local and global comprehension questions.

The results of the second research question revealed that vocabulary depth was a stronger predictor for TOEFL reading performance on Main Idea and Lexical Inferencing test items. As for the former test item, vocabulary depth made a strong

prediction of the HLCS, over and above the prediction made for the LLCS. In other words, the quality of vocabulary knowledge, which entails knowing the semantic association between and among words, sentences, and paragraphs, helped the test-takers grasp the main idea of texts with higher lexical cohesion. The contextual clues, therefore, assisted the learners in skimming and getting the gist of TOEFL passages, particularly the ones with higher lexical cohesion. Additionally, this aspect of vocabulary knowledge predicted the scores on the Lexical Inferencing test item of TOEFL reading subtest with high lexical cohesion, suggesting that contextual clues provided in high-lexical-cohesion reading test can substantially affect the contextual guessing ability of the test-takers. The results are in line with previous studies (e.g., Janebi Enayat & Babaii, 2018; Nassaji, 2006; Qian, 2002; Zhang & Anual, 2008). The elements of vocabulary depth are semantic associations among words like synonym and antonym as well as superordinate and subordinate words. For example, the words bird, animal, meat, egg, feather, wing, and eagle are all associated words. Research has shown that the utilization of semantic connections enhances lexical inferencing and the strategic utilization of such clues (Nassaji, 2004, 2006; Qian, 2004). Therefore, the level of vocabulary knowledge, considered as a variable of the reader, may impact the test takers' utilization of lexical chains and cohesion during text comprehension. The justification for this lies in the origins of these variables, both of which incorporate syntagmatic and paradigmatic relationships. The findings indicated that the extent of vocabulary knowledge could foresee the understanding of texts with strong lexical cohesion. Put differently, the semantic understanding of words linked to the depth of vocabulary may assist learners in identifying the lexical connections and utilizing them as a guide for cohesion. As noted by Nassaji (2004, 2006), individuals with greater vocabulary depth tended to perform better in lexical inferencing, with depth playing a crucial role in inferential success, a key aspect of reading comprehension.

The findings further revealed that vocabulary size was a stronger predictor variable for Lexical Inferencing test item of TOEFL LLCS. Vocabulary size refers to the quantity of vocabulary knowledge which does not consider the semantic associations between and among lexical items and/or segments of a text. This is in line with Alavi and Akbarian (2012) who found that vocabulary size was a strong

predictor for the lexical inferencing ability of the learners who took the TOEFL reading subtest. The students with higher vocabulary size and depth, as Nassaji (2004) argues, are more likely to succeed in using the available contextual cues in a text for guessing the meaning of unknown words. This follows a model developed by Meara and Wolter (2004) for the relationship between vocabulary size and depth in which knowing more words (vocabulary size) develops the knowledge of associations among words (vocabulary depth).

6. Conclusion

The findings of this study offer some implications for teachers, test designers, and materials writers. L2 teachers are advised to work on the students' vocabulary depth, as a reader variable which is more associated with text variables like lexical cohesion in text-dependent test performance. Developing the lexical network of the students should be prioritized as merely memorizing a list of words in a decontextualized manner would be useless for taking high-stakes exams like TOEFL reading. Test designers should pay more attention to lexical features of passages used for reading tests, particularly the TOEFL reading subtest. Such text features could interact significantly with the reader variables and affect the test-takers' performance. The development of high-stakes tests like TOEFL reading, therefore, requires detailed analysis of lexical features of texts. This suggests that traditional readability formulas cannot be reliable indicators of text difficulty, and lexical cohesion, for instance, should be considered in choosing suitable passages for the intended test-takers and/or learners. Materials writers should also reconsider the inclusion of various vocabulary exercises in coursebooks irrespective of the learners' vocabulary size and depth since the findings proved that only some task types were influenced by these dimensions of vocabulary knowledge. As proposed by Alavi and Akbarian (2012), the inquiry arises as to whether offering diverse forms of reading activities in course materials and classrooms, regardless of the vocabulary proficiency levels of language learners, is conducive to their comprehension and successful completion of tasks or materials, presuming that learners will achieve greater success after reaching a certain level of vocabulary proficiency. Moreover, this study offers some theoretical implications for the definitions of lexical cohesion and aspects of vocabulary knowledge. These terms

should not be viewed independently as both share similar meaning components, and their contributions should be probed as both reader and text variables.

Despite the potential implications of this study, future research endeavors must address a number of limitations to obtain more reliable and generalizable findings. Firstly, technological issues did not allow the administration of TOEFL iBT reading subtest. Although the reading skills covered in both modes of administration are similar, future studies could address this limitation to probe if the computer-based exam may yield different results. Secondly, the sample of this study was selected using convenience sampling which is restricted in generalizability of the results. Finally, the analysis of lexical cohesion could be conducted using indices from other computational tools like the Automatic Analysis of Cohesion (TAACO; Crossley et al., 2016b).

References

Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System*, *38*(3), 391–401. https://doi.org/10.1016/j.system. 2010.06.013

Alavi, S. M. H., & Akbarian, I. (2012). The role of vocabulary size in predicting performance on TOEFL reading item types. *System*, *40*(3), 376–385. https://doi.org/10.1016/j.system.2012.07.002

Alderson, J.C. (2000). *Assessing Reading*. Cambridge University Press.

Alharthi, T. (2020). Investigating the relationship between vocabulary knowledge and FL speaking performance. *International Journal of English Linguistics, 10*(1), 37–46. https://doi.org/10.5539/ijel.v10n1p37

Alshaar, R. (2008). *Measuring the stability of query term collocations and using it in document ranking* [Unpublished M.A. Thesis]. University of Waterloo.

Babaii, E., & Ansary, H. (2001). The C-test: a valid operationalization of reduced redundancy principle? *System*, *29*(2), 209–219. https://doi.org/10.1016/s0346-251x(01)00012-4

Babaii, E., & Fatahi-Majd, M. (2012). Failed restorations in the C-test: Types, sources, and implications for C-test processing. In Grotjahn, R. (Ed.), *Der C-test: Aktuelle Tendenzen/ The C-test: Current Trends*. Lang.

Babaii, E., & Moghaddam, M. J. (2006). On the interplay between test task difficulty and macro-level processing in the C-test. *System*, *34*(4), 586–600. https://doi.org/10.1016/j.system.2006.09.002

Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing*, *19*(4), 453–476. https://doi.org/10.1191/0265532202lt240oa

Bachman, L.F., & Palmer, A.S. (1996). *Language testing in practice*. Oxford University Press.

Çalışır, F., & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior*, *19*(2), 135–145. https://doi.org/10.1016/s0747-5632(02)00058-4

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, *18*(2), 119–135. https://doi.org/10.1016/j.jslw.2009.02.002

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016a). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, *32*, 1–16. https://doi.org/10.1016/j.jslw.2016.01.003

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016b). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Dabbagh, A., & Janebi Enayat, M. (2019). The role of vocabulary breadth and depth in predicting second language descriptive writing performance. *Language Learning Journal, 47*(5), 575–590. https://doi.org/10.1080/09571736.2017.1335765

Dabbagh, A., & Janebi Enayat, M. (2022). Comparing two measures of L2 depth of vocabulary knowledge using their relationship with vocabulary size. *Journal of Language and Education, 8*(3), 53–65. https://jle.hse.ru/article/view/13422

Derakhshan, A., & Janebi Enayat, M. (2020). High- and mid-frequency vocabulary size as predictors of Iranian university EFL students' speaking performance. *Iranian Journal of English for Academic Purposes, 9*(3), 1–13.

Enright, M. K., Grabe, W., Koda, K., Mosenthal, P., Mulcahy-Ernt, P. and Schedl, M. 2000: *TOEFL 2000 Reading Framework: a working paper*. TOEFL Monograph Series MS-17. Educational Testing Service.

Ferstl, E. C., & Von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cognitive Brain Research*, *11*(3), 325–340. https://doi.org/10.1016/s0926-6410(01)00007-6

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *UCLES*

Research Notes, 12, 8–10.

Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English.* Longman.

Hastings, A. (2002). Error analysis of an English C-test: Evidence for integrated processing. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Vol. 4, pp. 53–66). AKS-Verlag.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303–317. https://doi.org/10.1017/s0272263199002089

Henriksen, B. (2008). Declarative lexical knowledge. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Vocabulary and writing in a first and second language*. Palgrave Macmillan.

Hoey, M. (1991). *Patterns of Lexis in Text.* Oxford University Press.

Ishii, T., & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal, 40*(1), 5–22. https://doi.org/10.1177/0033688208101452

Jackson, N. E. (2005). Are university students' component reading skills related to their text comprehension and academic achievement? *Learning and Individual Differences*, *15*(2), 113–139. https://doi.org/10.1016/j.lindif.2004.11.001

Janebi Enayat, M., & Amirian, S. M. R. (2020). The relationship between vocabulary size and depth for Iranian EFL learners at different language proficiency levels. *Iranian Journal of Language Teaching Research, 8*(2), 97–114. https://doi.org/10.30466/IJLTR.2020.120891

Janebi Enayat, M., Amirian, S. M. R., Zareian, G., & Ghaniabadi, S. (2018). Reliable measure of written receptive vocabulary size: Using the L2 depth of vocabulary knowledge as a yardstick. *Sage Open, 8*(1), 1–15. https://doi.org/10.1177/2158244017752221

Janebi Enayat, M., & Babaii, E. (2018). Reliable predictors of reduced redundancy test performance: The interaction between lexical bonds and test takers' depth and breadth of vocabulary knowledge. *Language Testing, 35*(1), 121–144.

https://doi.org/10.1177/0265532216683223

Janebi Enayat, M., & Derakhshan, A. (2021). Vocabulary size and depth as predictors of second language speaking ability. *System*, *99*, 102521. https://doi.org/10.1016/j.system.2021.102521

Kaivanpanah, S., Akbarian, I., & Rezaee, M. (2022). The effects of reading and listening on L2 incidental learning and retention of different dimensions of word knowledge. *Language Related Research, 13*(5), 75–115. https://doi.org/10.52547/LRR.13.5.4

Kaivanpanah, S., & Alavi, S. M. H. (2008). The role of linguistic knowledge in word-meaning inferencing. *System*, *36*(2), 172–195. https://doi.org/10.1016/j.system.2007.10.006

Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, *4*(5). https://doi.org/10.4304/jltr.4.5.900-913

Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (2007) (Eds.). *Handbook of Latent Semantic Analysis.* Psychology Press.

MacMillan, F. (2007). The role of lexical cohesion in the assessment of EFL reading proficiency. *Arizona Working Papers in SLA & Teaching, 14*, 75–93.

Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, *72*, 23–36. https://doi.org/10.1016/j.system.2017.10.005

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System*, *52*, 1–13. https://doi.org/10.1016/j.system.2015.04.015

McDaniel, M. A., Hines, R. J., & Guynn, M. J. (2002). When text difficulty benefits Less-Skilled Readers. *Journal of Memory and Language*, *46*(3), 544–561. https://doi.org/10.1006/jmla.2001.2819

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts:

Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*(1), 51–62. https://doi.org/10.1037/h0087352

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*, 57–86. https://doi.org/10.1177/0741088309351547

McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Meara, P., & Wolter, B. (2004). V-Links: Beyond Vocabulary Depth. In Albrechtsen, D., Haastrup, K., & Henriksen, B. (Eds.), *Writing and Vocabulary in Foreign Language Acquisition* (pp. 85–96). Museum Tusculanum Press.

Mondria, J.A., & Wiersma, B. (2004). Receptive, productive, and receptive + productive L2 vocabulary learning: what deference does it make? In Bogaards, P., & Laufer, B. (Eds.), *Vocabulary in a Second Language* (pp. 79–100). John Benjamin Publishing Company.

Morris, J., & Hirst, G. (1991). Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational linguistics, 17*(1), 21–48.

Nassaji, H. (2004). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *Canadian Modern Language Review-revue Canadienne Des Langues Vivantes*, *61*(1), 107–134. https://doi.org/10.1353/cml.2004.0006

Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal*, *90*(3), 387–401. https://doi.org/10.1111/j.1540-4781.2006.00431.x

Nation, P. (2011). Research into practice: Vocabulary. *Language Teaching*, *44*(4), 529–539. https://doi.org/10.1017/s0261444811000267

Nguyen, B. T. T., Newton, J., & Tran, Q. P. N. (2022). The effect of topic on EFL writing by Vietnamese tertiary students: Insights from combining a lexical richness analysis with student self-reports. *Language Related Research, 13*(5), 511–540. https://doi.org/10.52547/LRR.13.5.19

Ozuru, Y., Dempsey, K. B., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003

Potelle, H., & Rouet, J. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, *58*(3), 327–345. https://doi.org/10.1016/s1071-5819(03)00016-8

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading Performance: An Assessment perspective. *Language Learning*, *52*(3), 513–536. https://doi.org/10.1111/1467-9922.00193

Qian, D. D. (2008). From single words to passages: Contextual effects on predictive power of vocabulary measures for assessing reading performance. *Language Assessment Quarterly*, *5*(1), 1–19. https://doi.org/10.1080/15434300701776138

Qian, D.D. (1998). *Depth of vocabulary knowledge: assessing its role in adult's reading comprehension in English as a second language* [*Unpublished PhD Dissertation*]. University of Toronto.

Qian, D.D. (2004). Second language lexical inferencing: preferences, perceptions, and practices. In Bogaards, P., & Laufer, B. (Eds.), *Vocabulary in a Second Language* (pp. 155–169). John Benjamins Publishing Company.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, *10*(3), 355–371. https://doi.org/10.1177/026553229301000308

Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, *18*(1), 1–32. https://doi.org/10.1177/026553220101800101

Sardinha, T.B. (2001). Lexical segments in text. In Scott, M., & Thompson, G. (Eds.), *Patterns of Text in honour of Michael Hoey* (pp. 213–238). John Benjamins Publishing Company.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, *64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., Ching Ng, J.W., & Garras, J. (2011). The Word Associates Format: Validation evidence. *Language Testing, 28*(1), 105–126. https://doi.org/10.1177/0265532210373605

Schoonen, R. & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25*(2), 211–236. https://doi.org/10.1177/0265532207086782

Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asian Social Science, 4* (12), 135–137.

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, *24*(1), 99–128. https://doi.org/10.1177/0265532207071513

Sigott, G. (2004). *Towards identifying the C-Test construct*. Peter Lang.

Silber, H., & McCoy, K. F. (2002). Efficiently computed lexical chains as an intermediate representation for automatic text summarization. *Computational Linguistics*, *28*(4), 487–496. https://doi.org/10.1162/089120102762671954

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139–152. https://doi.org/10.1080/09571730802389975

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Allyn and Bacon.

TOEFL ACTUAL TESTS. (2005). Ebteda Publications.

Tong, X., & Tong, X. (2022). How vocabulary breadth and depth influence bilingual reading comprehension: Direct and indirect pathways. *Learning and Individual Differences*, *100*, 102227. https://doi.org/10.1016/j.lindif.2022.102227

Uchihara, T., & Clenton, J. (2020). Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research, 24*(4), 540–556. https://doi.org/10.1177/1362168818799371

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. https://doi.org/10.1080/09571736.2016.119152

Webb, S. A., & Sasao, Y. (2013). New directions in vocabulary testing. *Language Testing, 44*(3), 263–277. https://doi.org/10.1177/0033688213500582

Webb, S., Sasao, Y., & Ballance, O. J. (2017). The updated vocabulary levels test. *ITL – International Journal of Applied Linguistics*, *168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Zhang, L. J., & Anual, S. B. (2008). The role of vocabulary in reading comprehension. *RELC Journal*, *39*(1), 51–76. https://doi.org/10.1177/0033688208091140

About the Authors

Mostafa Janebi Enayat is an Assistant Professor of Applied Linguistics at the University of Maragheh, Iran, where he teaches academic writing, language assessment, and computer-assisted language teaching. His main research interests include second language vocabulary and writing assessment, and educational technology. He has published in different international journals such as *System*, *Language Learning Journal*, *Review of Educational Research*, *Language Testing*, and *Innovation in Language Learning and Teaching*.

Farhad Mazlum is an Associate Professor at the University of Maragheh, Iran. He is interested in language planning and policymaking, English for academic purposes, and language assessment. He has published in *Language Policy, Journal of English for Academic Purposes, The Curriculum Journal, and International Journal of Educational Psychology*.