LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume: 16, No: 1, January – June 2023

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

The Role of Vocabulary Size and Depth in Predicting Postgraduate Students' Second Language Writing Performance

Apisak Sukying

apisak.s@msu.ac.th, Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand

APA Citation:
Sukying, A. (2023). The role of vocabulary size and depth in predicting postgraduate students' second language writing Performance. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 575-603.

Received 22/11/2022

Received in revised form 02/01/2023

Accepted 15/01/2023

ABSTRACT

This study investigated the contribution of vocabulary size and depth to second language (L2) writing performance. For this purpose, 53 English as a Foreign Language (EFL) postgraduate students took the Vocabulary Levels Test (VST) and the New Vocabulary Levels Test (NVLT), to measure vocabulary size, and the Word Associates Test (WAT) and the Productive Vocabulary Levels Test (PVLT), to assess vocabulary depth, as well as an argumentative writing task. The results of correlational and stepwise regression analyses indicated that the PVLT and the WAT significantly predicted the overall L2 argumentative writing performance, while both vocabulary size and depth were correlated to overall writing performance. Vocabulary depth was also predictive of the vocabulary component of L2 argumentative writing, and both vocabulary size and depth were correlated with the vocabulary component. Although vocabulary size and depth were essential for L2 writing performance, the current findings highlight the predictive role of vocabulary depth over and above vocabulary size in L2 argumentative writing. This is likely due to the criteria used to assess L2 writing and vocabulary, which emphasizes vocabulary diversity, complexity, and quality. The pedagogical

implications of these results and suggestions for future studies are also discussed.

Keywords: vocabulary size, vocabulary depth, L2 writing, argumentative writing

Introduction

Vocabulary knowledge is a critical challenge in language learning and plays a pivotal role in the overall language ability of second language (L2) learners (Cheng & Matthews, 2018; Janebi Enayat & Amirian, 2020; Koizumi & In'nami, 2013; Laufer, 2013; Milton, 2013; Nation, 2013; Qian & Lin, 2020; Sukying, 2018; Webb, 2020). Indeed, "without grammar, very little can be conveyed; without vocabulary, nothing at all can be conveyed" (Wilkins, 1972, p. 111). Learning vocabulary is a primary stage of learning a foreign language, and vocabulary is learned before grammatical rules. Vocabulary is also an essential component of mastering an L2 (Janebi Enayat & Derakhshan, 2021; Laufer, 2013; Nation, 2013, Qian & Lin, 2020; Schmitt, 2010). L2 learners are usually mindful that limitations in their vocabulary knowledge impact their communication skills as, without vocabulary, they are unable to express their thoughts (Nation, 2013). Schmitt (2008) noted that "one thing that students, teachers, materials writers, and researchers can all agree upon is that learning vocabulary is an essential part of mastering a second language" (p. 329). Alderson (2005) further argued that the learner's vocabulary size is related to their test performance, indicating that language ability is, to some extent, a function of vocabulary size. As such, measures of vocabulary knowledge have been developed and used as language placement tests (Janebi Enayat & Derakhashan, 2021; Milton, 2009; Nation, 2013; Schmitt, 2014).

Within L2 vocabulary acquisition, vocabulary knowledge has received significant theoretical and empirical attention (e.g., Schmitt, 2014; Webb, 2020). Due to its richness and complex nature, the construct of vocabulary knowledge has been theorized and conceptualized into different dimensions. One of the well-known dimensions is the size and depth of vocabulary knowledge (Henriksen, 1999; Read, 2000). Vocabulary size refers to the number of words that language learners know at a particular level of language proficiency (Nation, 2013). In most research, knowing a word has been synonymous with demonstrating knowledge about its form-meaning link, and it has been assumed that more frequent words are acquired earlier and better than less frequent ones (Laufer & Goldstein, 2004; Sukying, 2017, 2022). By contrast, vocabulary depth is the extent to which a vocabulary item is known or the quality of vocabulary knowledge and its lexical associations (Read,

2000). The operationalization of vocabulary depth is challenging but it is often conceived as knowledge of multiple components or dimensions.

The vocabulary size and depth dimensions have been widely studied in relation to their roles in L2 skills. Several studies highlight vocabulary size and depth as central determinants of L2 ability (Cheng & Matthews, 2018; Janebi Enayat & Amirian, 2020; Koizumi & In'nami, 2013; Laufer, 2013; Miralpeix & Muñoz, 2018; Qian & Lin, 2020; Read, 2004; Schmitt & Meara, 1997; Stæhr, 2009; Sukying, 2017, 2018). Regarding receptive skills, measures of vocabulary size and depth are significant indicators of listening comprehension (Cheng & Matthews, 2018; Ha, 2021; Matthews, 2018; Matthews & Cheng, 2015; Shen, 2008; Shiotsu & Weir, 2007; Stæhr, 2008, 2009; Wallace, 2020) and reading comprehension (Albrechtsen et al., 2008; Alavi & Akbarian, 2012; Cheng & Matthews, 2018; Ha, 2021; Qian, 1999, 2002; Zhang & Anual, 2008). Vocabulary size and depth also predict productive skills, especially speaking performance (Alharthi, 2020; Derakhshan & Janebi Enyarat, 2020; Janebi Enayat & Derakhshan, 2021; Koizumi & In'nami, 2013; Uchihara & Clenton, 2020; Uchihara & Saito, 2019) and writing ability (Atai & Dabbagh, 2010; Baba, 2009; Dabbagh & Janebi Enayat, 2019; Wu et al., 2019). Vocabulary size and depth can therefore determine the learner's L2 performance.

While several studies have examined the relationship between word knowledge aspects and receptive skills (reading and listening), far less research has been conducted on word knowledge aspects and productive skills (speaking and writing). Schmitt (2014) noted this paucity of research and stated that "it is an interesting, but unexplored, question whether the two would equally predict other kinds of language use" (p. 939). Moreover, previous studies investigating the relationship between vocabulary knowledge and L2 writing ability have focused on only one dimension: either size or depth of vocabulary knowledge. However, vocabulary knowledge requires more than knowledge of the form-meaning connection of the word (vocabulary size); it also requires the acquisition of multiple aspects of vocabulary knowledge (i.e., depth) (Henriksen, 1999; Meara & Wolter, 2004; Nation, 2013; Read, 2004; Schmitt, 2014).

Despite the increased interest in vocabulary size and depth as a construct for operationalizing the study of L2 skills, their role in EFL learners' writing ability has received little attention. Furthermore, research suggests that different tests measure different components of vocabulary knowledge (Read, 2000; Schmitt, 2010; Webb, 2013). This has led to insufficient empirical data on the multiple components measured by current vocabulary tests, and the relationship between these components and productive skills. The current study will, therefore, further investigate the relative contributions of size and depth of vocabulary knowledge to writing ability. The study also

seeks to determine whether vocabulary size and depth of knowledge predict EFL learners' L2 writing ability for argumentative writing essays. Understanding the role and contribution of the size and depth of vocabulary knowledge will shed light on the complex nature of vocabulary knowledge and its role in L2 ability and development in EFL learners.

Review of Literature

Theoretical framework

The construct of vocabulary knowledge, also referred to as word knowledge (e.g., Nation, 2013) or lexical knowledge (e.g., Laufer & Goldstein, 2004), is complex and multidimensional (Nation, 2013; Schmitt, 2010, 2014). This complex and multi-aspect construct has been conceptualized using a number of approaches. One approach is general proficiency with a word. This sustained conceptualization emerges from superficial recognition of a lexical item and gradually continues until the complete mastery of the word, defined as accurate language use (Read, 2000; Sukying, 2018; Wesche & Paribakht, 1996). A second approach to conceptualizing vocabulary knowledge is by deconstructing it into separate components, often described as a dimensional approach (Nation, 2013; Richards, 1976). This approach argues that different measures might be appropriate at the different stages of lexical acquisition. Vocabulary knowledge can also be defined as the automaticity with which the lexicon can be recognized and produced (Daller et al., 2007; Pellicer-Sánchez, 2019).

Henriksen (1999) also proposed three dimensions of lexical knowledge: a partial-to-precise knowledge dimension, a depth-of-knowledge dimension, and a receptive-productive dimension. The first dimension states that vocabulary knowledge is not an 'all-or-nothing' phenomenon (Laufer, 1998, p. 256). Indeed, vocabulary knowledge progresses from simple recognition of the lexical form and its meaning to varying levels of partial knowledge, and progresses to a fully mastered level of expertise. The development of partial-to-precise knowledge gradually moves and expands as a learner's exposure to the language increases (Henriksen, 1999). The second dimension highlights the continuum of knowledge depth and states that different test formats are required to assess the complete mastery of polysemous words. The receptive-productive continuum, also called passive and active knowledge (Laufer, 1998), defines the mastery of vocabulary knowledge in terms of the ability to control and use a specific word when speaking or writing. Laufer (2013) also argued that vocabulary incorporates form, lexical structure, syntactic pattern, meaning, associations, and collocations. For example, the lexical structures are referred to as the

morphological knowledge of a word, and lexical associations are reflected in the associations of the word with its synonyms, antonyms, and hyponyms. The syntactic pattern is related to the use of a word in phrases or sentences.

The distinction between size (breadth) and depth of vocabulary knowledge has been well established (Nation, 2013; Read, 2004; Schmitt, 2014; Webb, 2020). The size of vocabulary knowledge is defined as the number of words known by a language learner, at least partially. It has been argued that knowledge of form-meaning links is an initial proxy of vocabulary acquisition (González-Fernández & Schmitt, 2019; Laufer & Goldstein, 2004; Schmitt, 2008, 2010; Sukying, 2017, 2018; 2022; Sukying & Matwangsaeng, 2022; Sukying & Nontasee, 2022). In contrast, researchers have yet to reach a consensus on conceptualizing the depth of vocabulary knowledge (Milton, 2009; Nation, 2013; Read, 2004; Schmitt, 2010). Some argue that vocabulary depth entails how well a language learner knows an individual word, while others posit that the depth construct involves several aspects and components of vocabulary knowledge (Milton & Fitzpatrick, 2014; Nation, 2013; Schmitt, 2014; Schmitt & Schmitt, 2014). Depth of knowledge has also been defined as the different degrees to which lexical networks (or associations) are established from the perspective of polysemous meanings of words and collocations (Li & Kirby, 2015; Nation, 2013; Milton & Fitzpatrick, 2014; Read, 2000; Schmitt, 2014; Meara & Wolter, 2004). In this view, all these word knowledge aspects are seen as an incremental continuum of knowledge, which contribute to vocabulary depth (Milton & Fitzpatrick, 2014; Henriksen, 1999; Nation, 2013; Read, 2004; Schmitt, 2014). Overall, it appears that vocabulary depth refers to the quality of knowledge of the individual lexical item (Read, 2000; Schmitt, 2014).

The depth of vocabulary knowledge has also been defined as a lexical organization, which entails a number of links of related words in the lexical networks (Meara, 1997; Meara & Wolter, 2004; Schmitt, 2014). For example, the notion of *student* refers to a person who studies. This word might be related to other words, such as *school*, *teacher*, *class*, *learner*, *good*, *studious*, and *lazy*, but delineating how these words are learned and accumulated in a learner's mental lexicon, and how they are linked to other words, is not straightforward.

The contribution of vocabulary knowledge to L2 writing ability

The importance of vocabulary knowledge in writing has long been recognized (Coxhead, 2018; Dabbagh & Janebi Enyarat, 2019; Laufer & Goldstein, 2004; Li & Schmitt, 2009; Nation, 2013; Santo, 1988; Schmitt,

2010; Wu et al., 2019). Indeed, vocabulary knowledge plays a central role in L2 acquisition and contributes to overall language proficiency. One systematic inquiry (Alderson, 2005) revealed strong relationships between vocabulary knowledge and reading (.64), grammar (.64), listening (.61-65) and writing (.70-.79). Laufer and Goldstein (2004) also found that knowledge of the form-meaning link of words explained 42.6% of the total variance in students' overall English performance. These findings indicate that vocabulary size contributes significantly to overall language achievement.

More recently, Dabbagh and Janebi Enayat (2019) investigated the vocabulary size and depth of 67 English as a Foreign Language (EFL) students in Iran, using the Vocabulary Levels Test (VLT) as a vocabulary size measure and the Word Associates Test (WAT) to measure vocabulary depth in descriptive writing tasks. Their results indicated that vocabulary size was the only significant predictor of overall descriptive writing assessment, accounting for 19% of the variance in descriptive writing performance. Moreover, a correlational analysis revealed that vocabulary size and depth were both significantly associated with L2 writing performance. However, while significant relationships were found between the four word-frequency levels and descriptive writing scores, only the 5000 word-frequency level significantly predicted descriptive writing ability, accounting for about 18% of the variance in undergraduate students' writing ability. Their results align with earlier research by Laufer (2013), who showed that the lower word-frequency levels have a more robust predictive role in L2 writing ability.

Research has also shown that word choices are a strong predictor of the quality of written texts, and learners who used a variety of words appropriately and more low-frequency words obtained higher scores in their writing (Atai & Dabbagh, 2010; Derakhashan & Janebi Enayat, 2020; Engber, 1995; Li & Kirby, 2015; Wu et al., 2019). For example, Santos (1988) demonstrated that, according to advanced professors, lexical errors were the most problematic in EFL learners' writing. A more recent study by Li and Kirby (2015) also showed that vocabulary depth was closely and meaningfully associated with summary writing. They argued that the quality of lexical competence helps elicit word meaning in writing. Moreover, university students' writing typically progresses as their academic vocabulary increases (Albrechtsen et al., 2008). Dabbagh and Janebi Enayat (2019) further argued that vocabulary knowledge accounted for the most considerable amount of variation in writing composition.

Wu and her colleagues (2019) also found a strong relationship between EFL Chinese junior-high school students' vocabulary size and writing performance and indicated that vocabulary size made a greater contribution to writing tasks. Other studies also demonstrated strong associations between EFL learners' vocabulary size and writing performance

(Albrechtsen et al., 2008; Dabbagh & Janebi Enayat, 2019; Stæhr, 2008). However, it should be noted that some studies have shown that vocabulary size plays a negligible role in writing ability (Harrison et al., 2016; Schoonen et al., 2003, 2011). As such, there exists some inconsistency in the literature regarding the role of vocabulary size in EFL learners' writing ability.

Atai and Dabbagh (2010) used the Word Associates Test (WAT) as a vocabulary depth measure to study the role of in-depth knowledge of word use in a semantic network. They found that this knowledge significantly contributed to the overall writing ability of upper-intermediate EFL learners in Iran but not for their lower-intermediate peers. This result indicates that selecting appropriate lexical items with semantic precision can contribute to overall writing performance. This finding is consistent with a previous study by Baba (2009) showing that the semantic network of words and the ability to use this network contributes to writing ability.

Overall, previous studies have demonstrated that vocabulary size and depth separately contribute to EFL learners' writing ability. However, to the best of the researcher's knowledge, very few studies have examined the concurrent impact of vocabulary size and depth on EFL writing (Baba, 2009; Dabbagh & Janebi Enayat, 2019; Varnaseri & Farvardin, 2016; Wu et al., 2019). Moreover, the few existing studies have produced conflicting results. Some show that vocabulary size makes a more considerable contribution to writing performance than vocabulary depth (Dabbagh & Jane Enayat, 2019; Wu et al., 2019), while others have shown that depth of vocabulary knowledge is more important (Baba, 2009; Varnaseri & Farvardin, 2016). These inconsistent results may be due to the use of fairly limited measures of vocabulary size and depth to measure the learners' lexical proficiency.

Although vocabulary size and depth have been widely operationalized in L2 vocabulary research, compared to receptive skills, far less inquiry has been explored on their contribution to productive skills, especially writing. Vocabulary knowledge is essential to writing as L2 writers must be precise in order to help the readers understand the written text in detail. Indeed, L2 writers need to have a rich vocabulary size and a certain depth of knowledge to produce high-quality work. Therefore, the current study uses a test battery approach to investigate the relationship between vocabulary size and depth, and to determine whether these vocabulary knowledge aspects can predict the variation in L2 argumentative writing performance.

The current study

Considering the significant role of vocabulary size and depth in L2 writing ability, the current study aims to provide a clearer picture of the role of vocabulary size and depth in L2 writing ability. Specifically, it aims to

determine the extent to which vocabulary size and depth can predict postgraduate learners' writing performance. Four tests of receptive and productive vocabulary size and depth will be used, as well as the argumentative writing task.

In order to achieve these aims, the following research questions were formulated:

1. To what extent are vocabulary size and depth scores correlated with L2 argumentative writing performance?
2. What is the contribution of vocabulary size and depth scores in predicting the overall vocabulary component of an L2 argumentative writing task?

Methodology

Participants and context

The participants were 53 postgraduate students in an English Language Teaching (ELT) Program at a public university in Northeastern Thailand, where English was used as a medium of instruction. The justification for selecting the postgraduate students was that these participants were taught how to write an argumentation in English Thesis Writing (0105608), a required course for all postgraduate students in English Language Teaching Program at a university. The participants were recruited via convenience sampling and were informed that their decision on whether to participate or not would not affect their current or future relationship with the researcher or anyone else at the university. The participants were also told that they were free to withdraw from the study at any time.

The L1 for all participants was Thai, but all participants had studied English as a foreign language for more than 15 years at primary, secondary and tertiary levels of education. The Vocabulary Size Test (VST) was given to all participants to measure their vocabulary proficiency. The results indicated that the participants had a receptive vocabulary size of around 3,000-7,000 words, with an average of about 4,200 words. Their English language proficiency level was determined to be upper-intermediate based on the threshold score on the 2,000 high-frequency word band in the Vocabulary Levels Test (VLT) (Janebi Enayat et al., 2018).

Two PhD candidates in the English Language Teaching Program were trained to administer the tests and the writing task. Two university lecturers with at least 15 years of experience teaching writing in tertiary education were recruited to rate the writing task. Both were Thai native speakers who obtained their PhD degrees from a native English-speaking

country. Two training sessions were administered to train the raters to familiarize them with the scoring rubric.

Instruments

New Vocabulary Levels Test (NVLT)

The New Vocabulary Levels Test: Version 2 (Schmitt, Schmitt & Clapham, 2001) provides an estimate of vocabulary size at 2000, 3000, 5000 and 10000 frequency levels. The NVLT comprises six cue words and three different definitions, and each frequency level has 18 items. The test takers were required to match the cue words with their meaning. A correct definition match was awarded one point, and the maximum score was 72 points. Examples from the 5000-word levels are shown below:

1. area
2. contract _____ written agreement
3. definition _____ way of doing something
4. evidence _____ reason for believing something is or not true
5. method
6. role

This test was used because it is easy to manage, score and analyze. It is also widely used to assess vocabulary size and has a high level of reliability and validity (α=.88-.93) (e.g., Qian, 1999, 2002; Schmitt, 2014; Schmitt et al., 2001). In fact, many studies on vocabulary size and L2 language abilities in reception and production have employed the NVLT (e.g., Derakhshan & Janebi Enayat, 2020; Dabbagh & Janebi Enayat, 2019; Stæhr, 2009). It is also "perhaps the best known and most widely used vocabulary test" (Webb & Sasao, 2013, p. 264). Furthermore, the NVLT also minimizes the possibility of guessing since each set includes words from the same syntactic category. The test is freely available at www.lextutor.ca

Vocabulary Size Test (VST)

The Vocabulary Size Test (VST), developed and validated by Nation and Beglar (2007), is a measure of receptive vocabulary size and is used to assess knowledge of the form-meaning link of the words. The VST consists of 140 multiple-choice items from the 14 1,000 BNC word lists. Ten lexical items are included from each 1,000-word family band, and each represents a word family's members. The current study used the first ten frequency bands consisting of 1000 items. Each item probed the participants' knowledge of the form-meaning link of the words. The participants were asked to select the

most appropriate definition from the three distractors. The participants needed to have some idea of the word's meaning and form since the correct word and the distractors were in the same category and typically shared some elements of meaning. Below is a sample item from the VST:

0) **Jump**: She tries to **jump**.
   a. lie on top of the water
   b. get off the ground suddenly
   c. stop the car at the edge of the road
   d. move very fast

One point was awarded for the correct answer and no points were awarded for incorrect responses. The VST has high internal consistency reliability (Qian, 1999; Schmitt, 2014). It is widely used as it provides comprehensive, accurate, and reliable measures of the learners' written vocabulary size (e.g., Feng, 2016; Sukying, 2018). The VST is available online at www.lextutor.ca and www.vuw.ac.nz/lals/staff/paul-nation.

Productive Vocabulary Levels Test (PVLT)

The Productive Vocabulary Levels Test (PVLT), designed by Laufer and Nation (1999), measures the depth of vocabulary knowledge (vocabulary depth). The PVLT uses the gap-filling format presented in an unconstrained context where a sentence is provided, and the participants have to write the missing word. The current study incorporates four bands of frequency words, including the 2,000, the 3,000, the 5,000 and the 10,000. Each level consists of 18 items, yielding a total of 72 items. For each item, the first few letters of the cue words are provided to avoid non-targeted words that might fit semantically and grammatically in the sentence. For example, *The ar_____ of his office is 25 square meters*. Each item with the correct form and appropriate grammar received one point. Minor spelling mistakes were ignored, and no points were awarded for incorrect or blank responses. The justification for administering the PVLT in this study is that many studies have continually used it in vocabulary assessment and to examine the relationship between vocabulary knowledge and L2 skills (e.g., Alharthi, 2020; Schmitt & Meara, 1997; Sukying, 2018).

Word Associates Test (WAT)

The current study used the Word Associates Test (WAT) developed by Read (1993) to measure the depth of vocabulary knowledge. The WAT has been used in previous studies (e.g., Dabbagh & Janebi Enayat, 2019; Janebi Enayat & Derakhshan, 2021; Koizumi & In'nami, 2013; Schmitt, 2014; Schmitt et al., 2011), and its validity has been investigated (e.g., Schmitt et al.,

2011). The split-half reliability of the WAT was also analyzed, reaching 0.89 in Nassaji's (2006) study. The test contains 40 items, each of which provides an adjective word with a set of eight stimulus words. The test has two boxes on the left and right. The left box includes four possible synonymous words and four viable collocations of stimulus words of the stimulus on the right box. The WAT requires the participants to select the four related words or associates. The four selected words represent the concept of depth of knowledge through the selection of associates for a particular word. According to Read (1993, 2000), the target words are semantically related to associates in three ways: paradigmatic (semantic synonyms), syntagmatic (collocations) and analytic (dictionary definitions). The participants obtained one point for each correct response but were not penalized for incorrect answers. The maximum total score was 160. A sample question is shown below:

calm

open	quiet	smooth	tired	cloth	day	light	person

Writing task

An argumentative writing task was developed for this study. The participants were asked to write a five-paragraph argumentative essay of between 300 and 500 words in response to the given prompt. The argumentative essay is expected to be well-organized, well-supported, clear, concise, and coherent and should follow the key features of an argumentative essay. The participants were allowed to use a dictionary to support their writing and were given two hours to complete the task. The use of the dictionary would not affect participants' L2 writing performance because word spellings were not rated in their overall writing performance and vocabulary component scores.

The argumentative writing task was assessed on an analytic scale developed by Jacobs et al.'s (1981) model, which was slightly modified in order to suit the participants' context. Compared to the holistic scale, the analytic scale offers more detailed descriptions of the writer's performance on different writing features (Weigle, 2002) and provides a more reliable assessment scale (East, 2009). The scale used in the current study contains five aspects of writing: content, organization, language use, vocabulary, and mechanics (Dabbagh & Janebi Enayat, 2019; Weigle, 2002). Previous studies have demonstrated the high reliability and validity of this analytic scale (Connor-Linton & Polio, 2014; Dabbagh & Janebi Enayat, 2019).

Data collection

Before the data collection, all participants were informed of the research purposes. The size and depth of vocabulary knowledge measures were administered on two separate days. The depth measures were given to participants on the first day, and the vocabulary size tests were distributed on the second day. The depth of vocabulary knowledge measures was administered before the vocabulary size tests to prevent participants from drawing connections between the written forms and meaning of the words appearing on the tests of vocabulary depth. There was a 10-minute interval between tests to combat fatigue. Moreover, before the test battery was administered, the instructions and a few examples of the test were given to participants in their native language. The participants were allotted 60 minutes to complete each test. The argumentative writing task was administered a week after the vocabulary test battery, and participants were given two hours to complete it.

Data analysis

The study design included the size and depth of vocabulary knowledge as independent variables and the appraisal of participants' argumentative writing as the dependent variable. Descriptive statistics were calculated for the participants' test performance on the vocabulary test battery and the argumentative writing task. Inferential statistics were used to analyze this data. An independent-sample *t*-test was performed to compare the means of different score sets. This analysis was combined with a repeated-measures ANOVA to compare the variance between the means of related groups. This was achieved by stepwise multiple regression and a correlational analysis, which is a statistical means of arriving at an explanatory model for the variance. In this study, this analysis determined if vocabulary knowledge test performance could predict participants' L2 writing performance. The R^2 and associated statistics were also used to calculate the statistical significance of the overall regression model.

Assessing L2 writing essay

Two experienced university lecturers assessed the argumentative essay based on the analytic scale (Dabbagh & Janebi Enayat, 2019). First, they were trained to use the analytic scale using a sample of two essays. After the training session, the ratings were calibrated, and some guidelines were given to the assessors to ensure the most accurate assessment. Both assessors then analyzed all the argumentative essays. Moreover, to conform to the ethical

considerations for using the written tasks and to ensure a fair rating of the writing tasks, the personal information of the participants was not provided for the raters. A Pearson correlational analysis was conducted between the two assessors to determine the inter-rater reliability for each component and overall writing performance. Overall, the correlation coefficients indicated relatively high correlations between the two raters, as shown in Table 1.

Table 1

Pearson correlation coefficients between the two raters

Rater	C2	O2	V2	L2	M2	OW2
Content1	.892					
Organization1		.673				
Vocabulary1			7.33			
Language use1				7.30		
Mechanic1					.584	
Overall writing1						.816

Note: 1 = rater one; 2 = raters; $p < .001$

Findings

Descriptive statistics and correlation matrix

The descriptive statistics for the scores on the VST, NVLT, PVLT, WAT and the overall L2 argumentative writing performance are shown in Table 2. The vocabulary size performance consisted of the scores on the VST and NVLT, which were combined and calculated into percentages, while knowledge of vocabulary depth included the scores on the PVLT and WAT. As such, the combined scores would represent the overall vocabulary size and depth performance of Thai postgraduate participants. The descriptive statistics included the mean, standard deviation, minimum and maximum scores, and percentages of average performance of each test.

Table 2

Descriptive statistics for all tests and subtests (n =53)

Test	Total	Min.	Max.	Mean (%)	Std. deviation
VST	100	39	80	61.49 (61.49)	8.47
NVLT	120	43	110	77.85 (64.87)	11.97
Vocabulary size	220	100	190	139.34 (63.18)	19.14
PVLT	72	10	53	29.53 (41.01)	12.31

2000	18	7	18	13.08 (72.67)	3.01
3000	18	2	14	8.25 (45.84)	3.71
5000	18	0	14	5.68 (31.12)	4.08
10,000	18	0	9	2.53 (14.06)	2.89
WAT	160	30	121	91.49 (57.18)	24.92
Vocabulary depth	232	41	174	121.02 (49.10)	33.39
Argumentative writing	100	30	80	55.13 (55.13)	11.97
Vocabulary component	20	5	15	10.92 (54.60)	2.39

The results showed that, for vocabulary size, the participants performed better on the NVLT than the VST. For vocabulary depth, the participants achieved higher scores on the WAT than the PVLT. The students also achieved the highest scores on the 2000 frequency word level, followed by 3000 word-frequency level, 5000 and 10,000 word-frequency levels, indicating that students acquired vocabulary according to word frequency levels. This suggests that the participants acquired some aspects of vocabulary knowledge (vocabulary size) and gradually continued to develop their knowledge of vocabulary (depth of vocabulary knowledge). Together, the current findings suggest that learners acquire vocabulary knowledge at varying stages.

The results of the correlational analysis between vocabulary size and depth and overall L2 argumentative writing are shown in Table 3. All pairs of correlations were statistically significant and positive. The high correlation coefficients were found between vocabulary size and the NVLT (0.92), vocabulary depth and the WAT (0.95), and Vocabulary component and L2 writing (0.94). These values could indicate a lack of discriminant validity between the two constructs and suggest that there is an underlying component for all these tasks.

As previously mentioned, the NVLT is a receptive measure of vocabulary size that may assess the participants' ability to recognize the words that are orthographically and semantically related to the prompt words. However, the WAT is a productive measure of vocabulary depth that may capture the participants' ability to recall and retrieve the words that collocate with, or are semantically related to, the prompt words without being constrained on what words to use in language production (Read, 1993, 2000). The vocabulary component in the writing task is also related to the overall L2 writing performance. The higher performance on these tests could, therefore, help the participants to produce better writing with more complex words and structures.

Table 3

Pearson correlation coefficients among the vocabulary size and depth and L2 writing task and vocabulary use (n =53)

Test	1	2	3	4	5	6	7	8
VST (1)	-							
NVLT (2)	.47**	-						
PVLT (3)	.65**	.52**	-					
WAT (4)	.52**	.46**	.56**	-				
Vocabulary size (5)	.78**	.92**	.66**	.56**	-			
Vocabulary depth (6)	.63**	.54**	.79**	.95**	.66**	-		
L2 argumentative writing (7)	.42**	.40**	.54**	.50**	.47**	.57**	-	
Vocabulary component (8)	.39**	.33*	.47**	.53**	.41**	.57**	.94**	-

**$p < .001$; *$p < .01$ (two-tailed)

The contribution of vocabulary size and depth in predicting the overall L2 writing performance

As shown in Table 4, the results of regression analysis indicated that Model 1, where solely the PVLT was entered, accounted for about 29% of the variance in participants' argumentative writing performance ($R^2 = .291$, $F = 20.955$, $p < .001$), indicating that the PVLT significantly predicted participants' argumentative writing performance ($\beta = .540$, $t = 4.578$, $p < .001$). Likewise, Model 2, where the PVLT and the WAT were entered, explained ~35% of the variance in participants' argumentative writing performance ($R^2 = .347$, $F = 13.287$, $p < .05$). The inclusion of the WAT only added 6% to the prediction of variance achieved by the PVLT alone. In Model 2, both the PVLT ($\beta = .381$, $t = 2.769$, $p < .01$) and the WAT ($\beta = .285$, $t = 2.067$, $p < .05$). significantly predicted participants' argumentative writing performance. When the VST, the NVLT, the PVLT and the WAT were entered, as shown in Model 3, these variables accounted for ~36% of the variance in students' argumentative writing performance ($R^2 = .355$, $F = 6.597$, $p < .001$); however, no individual variables significantly predicted students' argumentative writing performance. Overall, the results of regression analysis indicated that knowledge of vocabulary depth predicted overall performance on the students' L2 writing task more strongly than knowledge of vocabulary size.

Note that vocabulary size was the combined scores of the VST and the NVLT, while the vocabulary depth consisted of the combined scores of the PVLT and the WAT. In Model 4 of the stepwise regression analysis, vocabulary depth accounted for 34.6% of the variance in overall L2 argumentative writing performance ($R^2 = .346$, $F = 26.991$, $p < .001$), and only

vocabulary depth significantly predicted participants' argumentative writing performance (β = .588, t = 5.195, p < .001). In Model 5 of Forced Entry to enter both vocabulary size and depth (Keith, 2014), the two predictors explained 35.4% of the variance in overall L2 argumentative writing performance (R^2 = .354, F = 13.680, p < .001), indicating that vocabulary size contributed about 1% to the overall writing score. However, only the vocabulary depth significantly predicted participants' argumentative writing performance (β = .502, t = 3.130, p < .01).

Table 4

Multiple regression for size, depth and L2 argumentative writing (n = 53)

Dependent	Model	Predictor	F	β	t	P	sr^2	R^2
L2 writing	1	(constant)	20.955		10.823	.000		.291***
		PVLT		.540***	4.578	.000	.540	
	2	(constant)	13.287		6.053	.000		.347*
		PVLT		.381**	2.769	.008	.365	
		WAT		.285*	2.067	.044	.281	
	3	(constant)	6.597		2.070	.044		.355***
		VST		.029	.180	.858	.026	
		NVLT		.099	.702	.486	.101	
		PVLT		.253	1.718	.092	.272	
		WAT		.328	1.958	.056	.241	
	4	(constant)	26.991		6.482	.000		.346***
		Vocabulary depth		.588***	5.195	.000	.588	
	5	(constant)	13.680		2.213	.031		.354***
		Vocabulary size		.123	.766	.447	.106	
		Vocabulary depth		.502**	3.130	.003	.405	

***p < .001; **p < .01; *p < .05

The contribution of vocabulary size and depth in predicting the vocabulary component

Stepwise multiple regression analyses were performed to determine the contribution that vocabulary size and depth might make to the vocabulary component in the L2 argumentative writing task. As illustrated in Table 5, two models emerged for the predictive power of vocabulary size and depth, and only the WAT was significant. In Model 1, in which the sole predictor variable was the WAT, 28% of the variance in participants' performance on the vocabulary component was explained (R^2 = .282, F = 20.079, p < .001).

The WAT significantly predicted the dependent variable in this model (β = .532, $t = 4.481$, $p < .001$). Model 2, where all variables were entered, accounted for ~33% of the variation in the vocabulary component of the argumentative writing task ($R^2 = .327$, $F = 5.824$, $p < .01$). Notably, only the WAT was a significant predictor in this model (β = .378, $t = 2.507$, $p < .05$). Overall, the results of stepwise multiple regression analyses showed that both models were significant and vocabulary depth performance, especially the WAT, tended to be more predictive of performance on the vocabulary component of the argumentative writing task than vocabulary size performance.

Table 5

Multiple regression for size, depth and vocabulary component of writing (n = 53)

Dependent	Model	Predictor	*F*	β	*t*	P	sr^2	R^2
Vocabulary component	1	(constant)	20.079		5.783	.000		.282***
		WAT		.532***	4.481	.000	.532	
	2	(constant)	5.824		2.141	.037		.327**
		VST		.036	.222	.826	.032	
		NVLT		.026	.180	.858	.026	
		WAT		.378*	2.507	.016	.340	
		PVLT		.221	1.292	.203	.183	
	3	(constant)	23.871		6.436	.000		.319***
		Vocabulary depth		.469***	4.886	.000	.565	
	4	(constant)	11.733		2.677	.010		.319**
		Vocabulary size		.034	.207	.837	.029	
		Vocabulary depth		.541**	3.287	.002	.422	

***$p < .001$; **$p < .01$; *$p < .05$

As shown in Table 5, the stepwise multiple regression analysis revealed that, in Model 3, only vocabulary depth was a significant predictor of the vocabulary component of overall argumentative writing performance (β = .469, $t = 4.886$, $p < .001$), explaining 31.9% of the variance in overall L2 argumentative writing performance ($R^2 = .319$, $F = 23.871$, $p < .001$). Model 4, where both vocabulary size and depth were entered, accounted for about 31.9% of the variance in participants' overall scores on vocabulary component ($R^2 = .319$, $F = 11.733$, $p < .01$); with only vocabulary depth significantly predicting participants' argumentative writing performance (β =

.541, $t = 3.287$, $p < .05$). This finding suggests that the vocabulary size has little impact on participants' overall scores on the vocabulary component of L2 writing. Overall, the current findings indicated that vocabulary depth predicted overall performance on the L2 argumentative writing task more robustly than vocabulary size.

Discussion

The current study focused on examining the relative contributions of vocabulary size and depth to L2 argumentative writing performance and the vocabulary component of writing tasks. Two measures of vocabulary size (the VST and the NVLT) and two productive vocabulary depth measures (the PVLT and the WAT) were used as the independent predictor variables.

The first research question examined the extent to which vocabulary size, as measured by the VST and the NVLT, and vocabulary depth, as measured by the PVLT and the WAT, could predict overall L2 argumentative writing performance. The results showed that combining the four measures of vocabulary size and depth could significantly predict Thai postgraduate students' overall L2 argumentative writing performance. However, as individual predictor variables, only vocabulary depth (measured by the PVLT and the WAT) could significantly predict the overall argumentative writing performance among Thai postgraduate students. The analysis of the results also reveals that the PVLT had a more significant impact on L2 argumentative writing task than the WAT, accounting for about 29% of participants' overall writing performance. These findings are in contrast with Dabbagh and Janebi Enayat (2019), who found that vocabulary size (breadth) alone could significantly predict Iranian EFL undergraduate learners' descriptive writing. Others have also reported that vocabulary size could significantly predict L2 speaking ability (Janebi Enayat & Derakhshan, 2021; Koisumi & In'nami, 2013).

However, the findings of the current study highlighted the predictive role of vocabulary depth in overall L2 argumentative writing performance. This is consistent with a previous study showing that vocabulary size and depth significantly contributed to writing scores (Varnaseri & Farvardin, 2016). Indeed, the current findings indicated a more robust predictive role for vocabulary depth in writing performance, whereas other studies found that vocabulary size was a stronger contributor to writing scores (Albrechtsen et al., 2008; Dabbagh & Janebi Enayat, 2019; Stæhr, 2008). The predictive power of vocabulary depth in argumentative writing performance could be explained by the scale of the scoring criteria used for assessing L2 argumentative writing, which requires assessors to draw attention to lexical sophistication,

linguistic and structural complexity and range of vocabulary. These requirements could have factored into the assessors' judgements to overestimate the function of vocabulary depth and underestimate the role of vocabulary size (Janebi Enayat & Dabbagh, 2021; Janebi Enayat, 2019). Indeed, previous findings showed that vocabulary size was a significant predictor in an overall descriptive writing task. This suggests that the widely used analytic assessment scale for writing that was used in the current study may not have allowed the assessors to focus on other aspects of vocabulary knowledge, such as collocations or word associations, that also reflect the quality of the writing (Baba, 2009; Janebi Enayat & Derakhshan, 2021; Laufer, 2013). In brief, the inconsistent findings could result from the different scoring criteria used for writing tasks (Dabbagh & Janebi Enayat, 2019; Varnaseri & Farvardin, 2016).

The tests themselves may have also impacted the extent to which vocabulary depth and size explain writing performance in the current study. The WAT was designed as a productive measure of vocabulary depth to assess the participants' ability to recall and retrieve the words that collocate with, or are semantically related to, the stimulus words (Read, 1993; 2000). Likewise, the PVLT measured the participant's ability to select not only the correct form but also appropriate grammar that might fit semantically and grammatically in the provided sentence. On the other hand, the receptive vocabulary size, measured through the VST and the NVLT, only measured the test takers' knowledge of the form-meaning link of vocabulary. Therefore, the test takers, who might only have some idea of the vocabulary's meaning and form, could select the possible word without knowing how to use it in genuine contexts.

Additionally, the present study indicated that both vocabulary size and depth were correlated with argumentative writing performance, but only vocabulary depth was a significant predictor for overall argumentative writing performance. This is partially in line with previous studies that vocabulary depth is associated with writing scores (Atai & Dabbagh, 2010; Baba, 2009; Dabbagh & Janebi Enayat, 2019). The current study provided evidence for a more robust correlation between vocabulary depth and argumentative writing, although both vocabulary size and depth showed moderate and large associations with argumentative writing performance, respectively. However, other studies argued that vocabulary size was a stronger predictor than vocabulary depth (Albrechtsen et al., 2008; Dabbagh & Janebi Enayat, 2019; Stæhr, 2008).

The results of the current study also reveal positive correlations between Thai postgraduate learners' knowledge of vocabulary size and depth and L2 argumentative writing performance. Specifically, the findings demonstrated a moderate positive association between vocabulary size

assessment and overall L2 argumentative writing performance, whereas large correlations were observed between vocabulary depth and argumentative writing performance (Cohen, 1988). These findings align with previous studies that L2 writing performance is more strongly correlated with vocabulary depth than vocabulary size (Varnaseri & Farvardin, 2016). Argumentative writing performance involves the ability to recall and retrieve morphological and semantic associations of words and produce the properly derived and inflected forms in writing, which may explain its strong association with vocabulary depth. Indeed, the current findings support the extant literature that vocabulary knowledge is linked to L2 writing ability (Albrechtsen et al., 2008; Atai & Dabbagh, 2010; Dabbagh & Janebi Enayat, 2019; Stæhr, 2008; Varnaseri & Farvardin, 2016). Moreover, the close interrelatedness between vocabulary size and depth and L2 argumentative writing performance lends evidence to support the claim that vocabulary knowledge is an essential element of L2 writing ability (Albrechtsen et al., 2008; Atai & Dabbagh, 2010; Dabbagh & Janebi Enayat, 2019).

The results also showed that both vocabulary size and depth contribute to the vocabulary component of the L2 argumentative writing performance. Specifically, the results indicated that the joint contribution of the four vocabulary size and depth measures was statistically significant. However, the WAT, a measure of productive vocabulary depth, was the only predictor variable that could uniquely account for the vocabulary component of the argumentative writing ability. The current findings are consistent with previous studies showing that the learners' scores on vocabulary depth and size were significantly linked to the vocabulary component of writing (Dabbagh & Janebi Enayat, 2019; Varnaseri & Farvardin, 2016). Consistent with Varnaseri and Farvardin (2016), the current findings also indicated that vocabulary depth, especially the WAT, was a stronger predictive contributor than vocabulary size in explaining the variance in vocabulary component scores of the argumentative writing task. These findings, however, contrast with previous studies showing that vocabulary size was a stronger predictor than vocabulary depth in writing performance (Albrechtsen et al., 2008; Dabbagh & Janebi Enayat, 2019; Stæhr, 2008). This is perhaps because the descriptors for assessing the vocabulary component require the assessors to emphasize lexical sophistication, academic and low-frequency words, and range of vocabulary. The descriptors did not encourage the assessors to pay attention to different aspects of word knowledge, such as word associations, collocations, and syntagmatic or paradigmatic relations to select word appropriateness in the scoring criteria for writing (Read, 2000; 2007).

Given the results of the current study and the discussion mentioned earlier, it can be concluded that various aspects of lexical knowledge need to be measured simultaneously to obtain a more complete picture of the

contributory role of vocabulary knowledge in L2 writing. The current study provides additional empirical evidence to highlight the essential role of vocabulary depth and vocabulary size in L2 ability (Albrechtsen et al., 2008; Baba, 2009; Dabbagh & Janebi Enayat, 2019; Ishii & Schmitt, 2009; Stæhr, 2008; Sukying, 2017; 2022; Sukying & Matwangsaeng, 2022).

Conclusion

The current study demonstrated that vocabulary size and depth were essential for L2 writing performance. Specifically, the current findings highlighted the predictive role of vocabulary depth over and above vocabulary size in L2 argumentative writing performance and the vocabulary component of writing. The present results also indicated that vocabulary size and depth correlated with overall argumentative writing performance and vocabulary component scores. However, only vocabulary depth could significantly predict EFL learners' argumentative writing performance and the vocabulary component of L2 writing. Together, the current findings provide empirical evidence that vocabulary size and depth play distinctive roles in L2 writing ability, and different measures are required to assess different aspects of lexical knowledge.

Implications and Suggestions for Future Studies

The current findings indicated that while vocabulary depth needs to be a priority in writing courses, improving the learner's ability to relate morphological and semantic links among lexical items deserves pedagogical attention and must be incorporated into L2 assessment criteria. Stakeholders of L2 instruction and assessment may wish to consider incorporating different aspects of word knowledge in the classroom. English language teachers and assessors should pay attention to designing a sufficient amount of both vocabulary size and depth tasks to assist students in enhancing their lexical knowledge and helping them to use lexical items in appropriate contexts. In assessing L2 writing and, in particular, the vocabulary component, an over-emphasis on the use of less frequent words may lead to skewed judgements of the quality of students' writing (Dabbagh & Janebi Enayat, 2019). As such, the current findings suggest that a broader view of what lexical knowledge involves should be adopted by both teachers and assessors. This view needs to include in-depth knowledge of words (vocabulary depth) and their use in writing tasks. Indeed, the current findings illustrate that vocabulary learning strategies may not be sufficient, and teaching declarative and procedural strategies may be necessary for improved writing performance. The current results also shed light on language testing

for diagnostic purposes, including how to administer the tests to observe students' strengths and weaknesses in vocabulary knowledge, and improving the design of language curricula, lesson plans and materials or tasks.

There are inevitably some limitations to this study that confine its generalizability. First, the number of participants was relatively small (53 Thai postgraduate students). Therefore, the current findings need to be replicated with larger and more diverse groups of participants. Second, the current study administered only the VST and the NVLT to measure participants' receptive vocabulary size, and the PVLT and the WAT to measure students' productive vocabulary depth. Administering different tests to measure the various aspects of vocabulary knowledge, as suggested by Read (2000) and Schmitt and Schmitt (2014), would yield additional information about the contribution of productive vocabulary depth to L2 writing skills. Moreover, the WAT is limited to measuring knowledge of collocation, polysemy, and synonyms. Future studies could examine how other aspects of vocabulary depth, such as affixes and derivatives, interact with L2 writing using newly developed tests or well-established tests like the Updated Vocabulary Levels Test (VLT) (Webb, Sasao, & Balance (2017) and the Word Part Levels Test (WPLT) (Sasao & Webb, 2017). Third, the topic for the argumentative essay writing may also limit the range of vocabulary and word selection for L2 writing. Experimental studies are also needed to determine whether the different amounts of training on writing skills can affect the relationship between vocabulary knowledge and writing performance. It is also necessary to examine whether the change in vocabulary size and depth mastery can affect its contribution to writing performance and vice versa. Finally, it is also suggested to investigate the relative contributions of receptive and productive vocabulary knowledge to other writing genres in both native and L2 speakers of English.

Acknowledgements

This research project was financially supported by the Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand.

About the Author

Apisak Sukying: An Assistant Professor of ELT and Applied Linguistics. He works at the Department of Western Languages and Linguistics, Mahasarakham University, Thailand. He completed his PhD in TESOL from the University of Sydney, Australia. His research interests include L2 vocabulary acquisition and development, SLA, Academic writing, ESP, and ELT methodology. His email is apisak.s@msu.ac.th

References

Alavi, S. M., & 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

Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Palgrave Macmillan.

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.

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

Atai, M.R., & Dabbagh, A. (2010). Investigating vocabulary depth and semantic set in EFL learners' vocabulary use in writing. *TELL Journal 4*(2), 27–49.

Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing*, *18*, 191-208. https://doi.org/10.1016/j.jslw.2009.05.003

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. https://doi.org/10.1177/0265532216676851

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum.

Connor-Linton, J., & Polio, C. (2014). Comparing perspectives on L2 writing: Multiple analysis of a common corpus. *Journal of Second Language Writing*, *26*(3), 1-9.

Coxhead, A. (2018). Vocabulary and second language writing. *The TESOL encyclopedia of English language teaching*, 1-6. https://doi.org/10.1002/9781118784235.eelt0533

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

Dabbagh, A. (2016). The predictive role of vocabulary knowledge in listening comprehension: depth or breadth? *International Journal of English Language and Translation Studies*, *4*(3) 1–13.

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

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.

East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing Writing*, *14*(2), 88-115.

Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL composition. *Journal of Second Language Writing*, *4*(2), 139-155.

González-Fernández, B., & Schmitt, N. (2019). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics, 41*(4), 481-505. https://doi.org/10.1093/applin/amy057

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K., & Spurling, J. (2016). Predictors of spelling and writing skills in first-and second-language learners. *Reading and Writing*, *29*(1), 69-89.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303-317. https://doi.org/10.1017/S0272263199002089

Ha, H. T. (2021). Exploring the relationships between various dimensions of receptive vocabulary knowledge and L2 listening and reading comprehension. *Language Testing in Asia*, *11*(1), 1-20. https://doi.org/10.1186/s40468-021-00131-8

In'nami, Y., & Koizumi, R. (2016). Task and rater effects in L2 speaking and writing: A synthesis of generalizability studies. *Language Testing*, *33*(3), 341-366.

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

Jacobs, H., Zinkgraf, S., Wormuth, D., Harfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Newbury House.

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., & 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

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*, 900-913. https://doi.org/10.4304/jltr.4.5.900-913

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*(2), 255-271. https://doi.org/10.1093/applin/19.2.255

Laufer, B. (2013). Vocabulary and writing. In C.A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (1-5). Wiley-Blackwell.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*(3), 399-436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*, 33-51. https://doi.org/10.1177/026553229901600103

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics, 36*, 611-34. https://doi.org/10.1093/applin/amu007

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, *18*, 85-102.

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.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In Schmitt, N. & McCarthy, M. (eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 35-53). Cambridge University Press.

Meara, P., & Wolter, B. (2004). V-Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Writing and vocabulary in foreign language acquisition* (pp. 85-96). Museum Tusculanum Press.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindquist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use*: *New perspectives on assessment and corpus analysis* (pp. 57-78). EUROSLA Monographs.

Milton, J., & Fitzpatrick, T. (2014). *Dimensions of vocabulary knowledge*. Palgrave Macmillan. https://doi.org/10.1017/S0272263114000382

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 https://doi.org/10.1515/iral-2017-0016

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.

Pellicer-Sánchez, A. (2019). Examining second language vocabulary growth: Replications of Schmitt (1998) and Webb & Chang (2012). *Language Teaching*, *52*(4), 512-523. https://doi.org/10.1017/s026144481800037X

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, *56*(2), 282-307. https://doi.org/10.3138/cmlr.56.2.282

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., & Lin, L. H. F. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies*, (pp. 66–80). Routledge.

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. (2000). *Assessing vocabulary*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732942

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, *24*, 146-161. https://doi.org/10.1017/S0267190504000078

Richards, J.C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, *10*(1), 77-89.

Santos, T. (1988). Professors' reactions to the academic writing of nonnative-speaking students. *TESOL Quarterly*, *22*(1), 69-90.

Sasao, Y., & Webb, S. (2017). The Word Part Levels Test. *Language Teaching Research, 21*(1), 12-30. https://doi.org/10.1177/1362168815586083

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual.* Palgrave Macmillan. http://dx.doi.org/10.1057/9780230293977

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

Schmitt, N., & Meara, P. (1997). Research vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition, 19*(1), 17-36. https://doi.org/10.1017/S0272263197001022

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, *47*(4), 484-503. https://doi.org/10.1017/S0261444812000018

Schmitt, N. Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing, 18*(1), 55-88. https://doi.org/10.1177/026553220101800103

Schoonen, R., Gelderen, A. V., Glopper, K. D., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning*, *53*(1), 165-202.

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.

Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asian Social Science*, *4*(12), 135-137. https://doi.org/10.5539/ass.v4n12p135

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

Sukying, A. (2017). The relationship between receptive and productive affix knowledge and vocabulary size in an EFL Context. [Doctoral dissertation, University of Sydney, Australia].

Sukying, A. (2018). Investigating receptive and productive affix knowledge in EFL learners. In D. Hirsh (Ed.), *Explorations in second language vocabulary research* (pp. 183-218). Peter Lang.

Sukying, A. (2020). Word knowledge through morphological awareness in EFL students. *TESOL International Journal, 15*(1), 74-85.

Sukying, A. (2022). A taxonomy of English affix acquisition in EFL learners. In D. Hirsh (Ed.), *Research perspectives in language and education* (pp. 49-82). Peter Lang. https://doi.org/10.3726/b19346

Sukying, A., & Matwangsaeng, R. (2022). Exploring primary school students' morphological awareness in Thailand. *World Journal of English Language*, *12*(6), 388-401. https://doi.org/10.5430/wjel.v12n6p388

Sukying, A., & Nontasee, W. (2022). The acquisition order of vocabulary knowledge aspects in Thai EFL students. *World Journal of English Language, 12*(5), 306-319. https://doi.org/10.5430/wjel.v12n5p306

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

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*(4), 577–607 https://doi.org/10.1017/s0272263109990039.

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. *Language Learning Journal*, *47*(1), 64-75. https://doi.org/10.1080/09571736.2016.1191527

Varnaseri, M., & Farvardin, M. (2016). The relationship between depth and breadth of vocabulary knowledge and writing performance of Iranian MA students of TEFL. *Modern Journal of Language Teaching Methods*, *6*(2), 544-554.

Wallace, M. (2020). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning*, 58, 1-10. https://doi.org/10.1111/lang.12424

Webb, S. (2012). Depth of vocabulary knowledge. *The Encyclopedia of Applied Linguistics*. https://doi.org/10.1002/9781405198431.WBEAL1325

Webb, S. (2013). Depth of vocabulary knowledge. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–8). Blackwell Publishing Ltd.

Webb, S. (Ed.). (2020). *The Routledge handbook of vocabulary Studies* (Vol. 2). Routledge.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, *168*(1), 33-69. https://doi.org/10.1075/itl.168.1.02web

Wu, S., Dixon, Q., Sun H., & Zhang, P. (2019). Breadth or depth: The role of vocabulary in Chinese English-Language beginning writers' development, *International Journal of Bilingual Education and Bilingualism*, 1-17. https://doi.org/10.1080/13670050.2019.1572066

Wilkins, D. (1972). *Linguistics in language teaching*. Arnold

Weigle, S.C. (2002). *Assessing writing*. Cambridge University Press.

Wesche, M., & Paribakht, T. S. (1996). Assessing secondary language vocabulary knowledge: depth versus breadth. *The Canadian Modern Language Review, 53*(1), 13-40. https://doi.org/10.3138/cmlr.53.1.13

Zareva, A. 2005. Models of lexical knowledge assessment of second language learners of English at higher levels of language proficiency. *System*, *33*(4), 547–62.

Zhang, B., & Anual, S.B. (2008). The role of vocabulary in reading comprehension: the case of secondary school students learning English in Singapore. *RELC Journal*, *39*(1), 51–76.