

Investigating the Role of Word Knowledge Components in Chinese L2 Writing Ability

CHANGYONG MIN

Department of Western Language and Linguistics
Mahasarakham University, Thailand

APISAK SUKYING *

Department of Western Language and Linguistics
Mahasarakham University, Thailand
apisak.s@msu.ac.th

ABSTRACT

A body of research has looked into the nature of multiple-word knowledge components in recent years. However, the individual role of these components in L2 writing proficiency still remains unclear. The present study examined the interrelations between six-word knowledge components and explored the relationships between the lexical components and productive word use and L2 writing ability. The study adopted a multi-task approach using the word knowledge framework by Nation (2013, 2022) and the model of word knowledge components required in writing proposed by Coxhead (2007). Six measures, including one receptive word component (overall word size) and five productive depth knowledge components (productive form and meaning, association, productive derivation and collocation), were designed and validated to measure 147 Chinese EFL university learners' word knowledge relative to their word use and argumentative writing ability. The correlation and regression results demonstrated that derivative form production was the best predictor of word use and L2 writing ability. Its contribution was even stronger than productive form and meaning, though the latter two were also closely related to L2 writing and word use. Association and collocation predicted less variance yet still correlated with productive skills. However, the receptive size measured by the VLT had no correlations with L2 writing and lexical proficiency. Overall, this study provides empirical evidence for the theoretical word knowledge models and yields nuanced ideas regarding the smallest lexical predictors of L2 writing. Pedagogical implications for EFL vocabulary pedagogy are also discussed.

Keywords: Chinese EFL university learners; L2 writing ability; multiple word tests; word knowledge components; word size and depth

INTRODUCTION

Vocabulary knowledge can be construed as the knowledge of multiple-word components (Nation, 2022). These components are manifold, including word form, meaning, collocation, frequency of use, and association, among other aspects ((Milton & Fitzpatrick, 2014; Nation, 2022). Research has shown that the various lexical components are closely related to English as a Foreign Language (EFL) learners' lexical abilities and overall language proficiency (Choo et al., 2017; Qian & Lin, 2020), particularly second language (L2) writing ability (Kim et al., 2022; S. Wu et al., 2019). Thus, multiple tasks might be required to capture each component of vocabulary knowledge in L2 writing proficiency. Such tasks might help to extract detailed insights into the role of word knowledge components in an overall L2 argumentative writing task and actual vocabulary use (Yanagisawa & Webb, 2020).

Previous research has mainly focused on the dominant role of form-meaning links in L2 writing ability (Stæhr, 2008). Indeed, form and meaning are the most fundamental components in any kind of language production. This has been demonstrated in Chinese EFL sentence writing (Zhong, 2016), L2 junior-high students' free writing (S. Wu et al., 2019), and self-reported argumentative writing performance (Coxhead, 2007). However, the vocabulary-writing relation not only depends on the form-meaning connections (i.e., vocabulary size); it also requires the acquisition of multiple shades of meaning, semantic associations, grammatical functions and constraints on use in the writing process (i.e., vocabulary depth) (Read & Dang, 2022). Beyond vocabulary size, word depth components have also shown predictive relevance to L2 writing proficiency, such as academic register (Coxhead, 2007), accurate collocation use (Bestgen, 2017; Crossley et al., 2015) and derivational knowledge (Leontjev et al., 2016). Therefore, general word knowledge (overall word size) and specific word knowledge (depth components) should be combined to determine individual and group differences in vocabulary knowledge in L2 written production.

As such, research has called for multiple, separate tests to measure the individual role of word knowledge components in context (Schmitt, 2019). However, the few, if any, studies taking the challenge to assess multiple word components in writing simply captured a constrained range of them, such as receptive aspects only (Zhong, 2016) or merely one or two depth components (Bestgen, 2017; Leontjev et al., 2016; Sukying, 2023). As a result, studies measuring a wide range of word components concurrently associated with L2 writing are still needed, and the role of multiple word components in L2 writing ability remains unclear. The present study, therefore, focused on productive word depth components in a Chinese argumentative writing context and attempts to quantitatively explore the correlations and contributions of these components to L2 writing and word use. Two research questions will guide the study:

1. What are the correlations among the six-word components, namely, vocabulary size, productive form and meaning, association, productive derivation and collocation, and their relationship with Chinese EFL learners' writing score and word use?
2. To what extent do these discrete word knowledge components contribute to Chinese EFL learners' word use and overall L2 writing ability?

LITERATURE REVIEW

VOCABULARY SIZE AND DEPTH COMPONENTS

Vocabulary size/breadth-depth distinction has been widely acknowledged to describe the overall state of vocabulary knowledge (Li & Kirby, 2015; McLean et al., 2020; Read & Dang, 2022). Size indicates the number of words learners can identify, while depth points to the richness of word knowledge or how well learners know the words (Read, 2004). Size can be relatively straightforward to conceptualise in assessments as it is closely connected to aspects of word form and meaning (Schmitt, 2014). Depth is more complicated than size since it goes beyond form and meaning connections and involves the retrieval and use of “a rich and specific meaning representation as well as knowledge of the word's format features, syntactic functioning, collocational possibilities, register characteristics, and so on” (Read, 2004, p.155).

Following this line, research has typically addressed the depth construct from the component perspective (Nation, 2022; Schmitt, 2014). Nation (2013, 2022) provided the most comprehensive framework of word knowledge components, mainly featuring word **form** (pronunciation, spelling and word parts), **meaning** (form-meaning links, concepts referents and associations), and **use** (collocations, grammatical functions and constraints on use) (see Table 1). These word components can all be acquired at both receptive and productive levels. Following Nation’s (2022) definition, receptive word knowledge in the present study refers to the recognition of form-meaning association, while productive word knowledge involves the recall and use of multiple word depth components in sentence production and L2 writing.

TABLE 1. What is involved in knowing a word (Nation, 2013, p. 49)

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
P		How is the word written and spelt?	
Meaning	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
associations	R	What other words does this make us think of?	
	P	What other words could we use instead of this one?	
Use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this?
	constraints on use	R	Where, when and how often would we expect to meet this word?
		P	Where, when and how often can we use this word?

The component approach to vocabulary knowledge has fostered research into the nature of acquiring and using a word. For example, research has parsed the grammatical knowledge of word parts, i.e., inflectional and derivative affix knowledge (Schmitt & Zimmerman, 2002) and receptive and productive mastery of multiple word components (González-Fernández & Schmitt, 2020). These researchers demonstrated that the measured word components are interrelated, and some components may be easier to master than others. It is also suggested that multi-component vocabulary tests, though time-consuming and limited in word number (Read, 2004), are worthwhile since “such research produces a very detailed description of vocabulary knowledge, making it well worth the effort” (Schmitt, 1998, p. 286).

As such, Schmitt (2014) and Milton and Fitzpatrick (2014) encouraged the measurement of the various word components concurrently at both receptive and productive levels with a battery of tests. Inspired by this, González-Fernández and Schmitt (2020) employed extensive receptive and productive tests to examine the recognition and recall knowledge of four vocabulary components: form-meaning links, derivational forms, polysemy and collocation. More recently, Read and Dang (2022) focused on various knowledge components of academic vocabulary. Beyond form-meaning links, their study devised newly developed depth tests to measure synonyms, collocations and word parts. One major finding from these studies confirmed that the various word components contribute to the overall word knowledge construct at different levels. Another finding is that the acquisition order of these components varies markedly. For example, EFL learners have the best mastery of synonyms, followed by word parts and collocation (Read &

Dang, 2022) and word recall usually lags behind recognition (González-Fernández & Schmitt, 2020). While ample research has been done on the nature of word knowledge components, limited studies have captured these components related to actual language proficiency, such as L2 writing ability and word use. Nation (2022) also pointed out that his convenient description of vocabulary knowledge cannot explicate how these hypothesised components behave in actual use and that the relationship between word knowledge and writing is the least known to date.

VOCABULARY KNOWLEDGE COMPONENTS AND L2 WRITING

Building on Nation’s (2013, 2022) comprehensive framework of vocabulary knowledge, Coxhead (2007) specified what components of a word need to be acquired in producing a piece of an essay (See Table 2). Coxhead (2007) conducted empirical studies under the multi-component model and suggested that word components such as collocation, association, and register are critical factors in using a word apart from the basic form and meaning knowledge. L2 university learners reported that they had to mobilise knowledge of various lexical components during the writing process. Nation (2022) suggested that it is relevant to look into how the various components of a word fit into the language production process. Based on the lexical frameworks of Nation (2013, 2022) and Coxhead (2007), the present study probed into the relationship between different word knowledge components and L2 writing proficiency.

TABLE 2. Knowledge is required for the production of a word in writing (Coxhead, 2007, p. 332)

Form		How is the word written and spelt?
Meaning	Form and meaning	What word form can be used to express meaning?
	Concepts and referents	What item can this concept refer to?
	Association	What other words can we use instead of this one?
Use	Grammatical function	In what patterns must we use this word?
	Collocations	What words or types of words must we use with it?
	Constraints of use	Where, when and how often can we use this word?

These word knowledge components indeed figure prominently in L2 writing because they can make or break written communication (Qian & Lin, 2020). More specifically, the various types of word knowledge can be a major driving force for grammatical structures, which has been described as a lexical-driven procedure. In this process, vocabulary serves as a critical mediator between the representations and grammatical encoding. That is, a sentence can be produced only when there is a range of proper lexical knowledge behind the triggering message (ideas) and grammatical properties. Thus, the chosen word determines a sentence’s grammar, morphology and phonology (Nation, 2022).

Also, under the frameworks of Nation (2022) and Coxhead (2007), Zhong (2016) investigated the relationships between five receptive word components in L2 sentence writing tasks and identified that receptive meaning and form explained a 74.1% variance in sentence writing performance. Word class, association, and collocation also significantly improved the regression models, supporting Coxhead’s (2007) theoretical model. Other multi-component studies in L2 writing (Dabbagh & Janebi Enayat, 2019; S. Wu et al., 2019) also demonstrated that form and meaning knowledge can be strong predictors of L2 writing scores. Dabbagh and Janebi Enayat (2019) compared the size of the Vocabulary Levels Test (VLT) and the depth of the Word Association Format (WAF) and found that the former better predicted descriptive writing scores. S. Wu et al. (2019) also identified that vocabulary size contributed more to L2 writing quality than

depth components, such as adjective synonyms and morphological awareness. These results are consistent with previous studies documenting that vocabulary size robustly contributes to L2 writing (Stæhr, 2008). However, only receptive word knowledge (Stæhr, 2008; Zhong, 2016) or limited depth aspects (Dabbagh & Janebi Enayat, 2019) were touched on in previous studies, ignoring many other receptive and productive depth components.

Indeed, word depth components have also been demonstrated to relate closely to L2 writing proficiency. Previous research has explicitly focused on collocation and derivative knowledge in L2 writing performance. One of these studies was done by Bestgen (2017), who empirically showed that collocation is more critical than word diversity and sophistication in ESL writing ability. The same finding was obtained by Crossley et al. (2015), who found that the accurate use of collocations is the best predictor of L2 writing output, explaining 84% of the variance. Collocation, therefore, should be paid special attention to since L2 writers at advanced levels still find it difficult to use collocations correctly in their writing (Laufer & Waldman, 2011). In addition, Leontjev et al. (2016) revealed strong correlations between different types of derivational forms and L2 writing ability. These results suggest that vocabulary size, characterised by word diversity, sophistication or frequency, is less predictive of L2 writing ability than depth components. Yet, despite these fruitful results, researchers are still struggling to reach a consensus on the predictive power of the lexical components on writing ability. For example, collocation, association and academic register have their roles in writing scores in some studies (Coxhead, 2007; Crossley et al., 2015) but cannot predict writing ability in other studies (Waluyo & Bakoko, 2021; S. Wu et al., 2019; Zhong, 2016). This provides a rationale for the present study to holistically investigate how these lexical components behave in L2 writing.

The various components of vocabulary size and depth knowledge indeed bear a close relationship to L2 writing proficiency, be it academic vocabulary (Sukying, 2023) or low-frequency vocabulary (Dabbagh & Janebi Enayat, 2019). However, while a large amount of research attention has been paid to word knowledge components and reading comprehension ($r = .57 - .95$, Li & Kirby, 2015; McLean et al., 2020), the critical role played by different types of word knowledge in L2 writing has been relatively unattended (Qian & Lin, 2020).

METHODOLOGY

PARTICIPANTS

This study adopted a cross-sectional design to investigate 147 third-year Chinese university EFL learners who were English Education majors. The participants comprised 135 females and 12 males, aged between 19 and 21 years old. The gender difference represents the uneven proportion of the whole male and female populations of English learners in the university. Before the tests, the participants had a basic knowledge of argumentative writing as they all obtained credits in the writing courses. In the given semester, they were enrolled in eight courses related to the English language, making up approximately 14 hours of exposure to English every week. They all had passed the Test for English Majors Band 4 (TEM4), which is mainly designed to capture vocabulary and grammar, representing an intermediate English proficiency level of Chinese EFL learners. The participants' vocabulary level was estimated by the scores of the 2000 (96.5%), 3000 (92.1%), 5000 (76%) and 10000 (30.46%) sections of the Vocabulary Levels Test (Schmitt et al., 2001), with a compound score of 73.5% at the four frequency bands.

MATERIALS AND INSTRUMENTS

SELECTING TARGET WORDS

Twenty target words, as shown in Table 3, were sampled from the participants' textbook *Advanced English: Reading and Writing Course*. These words were also academic words appearing on the Academic Word List (AWL) (Coxhead, 2000). The AWL provides guidance on the most useful and valuable words for university students with academic writing goals.

TABLE 3. The twenty target words for depth tests

Word class	Numbers	Target words
Verbs	10	achieve, consume, detect, expand, indicate, justify, rely, restrain, stimulate, persist
Nouns	5	approach, access, lecture, intelligence, schedule
Adjectives	5	devoted, exposed, inclined, prior, voluntary

The selection criteria were designed to measure all word knowledge components: productive form and meaning, association, derivation and collocation. The sample included ten verbs, five nouns and five adjectives, as approximately 50% of the AWL words are verbs (Coxhead, 2000). Some words, such as *access* and *approach*, have two-word classes, which would not cause confusion as there was a prompt for the measured word class before each sentence in the derivational test. Because participants had to randomly use the target words in their L2 writing tasks, these words needed to fit in with the contexts of the L2 writing topics. In order to control the difficulty of the target words, two words were selected from each sublist of the AWL, yielding a total of twenty.

PRODUCTIVE DEPTH TESTS

There were five sections in the productive depth tests, with each section separately targeting each productive word knowledge component: productive form, productive meaning (word pair test), derivative test, association and collocation production. All the five depth tests used the same set of twenty target words to elicit different dimensions of lexical knowledge (Schmitt, 2019; Yanagisawa & Webb, 2020).

Productive form was captured by the Productive Vocabulary Levels Test (PVL) developed by Laufer and Nation (1999) (see the Appendix for examples of instruments). Laufer and Nation (1999) suggested that the PVL is related to the ability to produce words in writing. It is noted that the adapted PVL in the present study was not frequency-based and was solely limited to the twenty target words. Since the PVL is a productive test, the minimum number of letters was prompted as long as test-takers could eliminate the possibility of choosing alternative words.

Productive meaning was measured by the active (productive) recall test in Laufer and Goldstein (2004). Empirical studies have shown that a productive L1-L2 word pair can be a more practical approach to vocabulary knowledge if the learning goal is productive use (Nation, 2022). The first letter of the target word was provided as a prompt to prevent participants from using other words with the same meaning. All the Chinese word translations were selected from the definitions of the target words in the *Oxford Advanced Learner's English-Chinese Dictionary, ninth edition*.

The productive derivative test used the test format developed by González-Fernández and Schmitt (2020). The original format required participants to fill in the gaps in each of the four semantically similar sentences using the appropriate derivational forms of the target words. The

present study adapted this test, only eliciting three parts of speech (verb, noun and adjective) for one target word. Adverbs were excluded since most adjective forms have already signalled adverb knowledge. Learners are likely to have adverb knowledge if they can produce the adjective form of a word (Zhong, 2016).

The WAF created by Read (2004) was modified to measure participants' synonymy knowledge. Since the original WAF is context-independent (Read, 2004) and difficult to score (Schmitt, 2014), the present study split the WAF into a single synonymy test in sentence contexts. There were two to three keys in the box to minimise the possibility of correct guessing. The associates were selected based on similar core meanings, which could replace the target words in the sentences. All the synonyms of the target words were sourced from the thesauruses of the *Longman Dictionary of Contemporary English* and *Merriam-Webster Dictionary*.

Collocation production was adapted from González-Fernández and Schmitt (2020), who took a form-recall approach to measure productive collocation knowledge in sentence contexts. The first letters of the words in the collocations were given in the sentence gap, and a related L1 sentence was prompted as the contextual information. The measured collocations were sampled from the phrase banks in the *Longman Dictionary* and *Macmillan Dictionary*.

It is noted that all the sample sentences in the depth tests were selected from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). Sentences with unfamiliar contexts and difficult words were excluded.

RECEPTIVE SIZE TEST

The VLT refined by Schmitt et al. (2001) was used to capture the receptive size of vocabulary knowledge since this test has been universally accepted as “the closest thing we have to measure vocabulary knowledge” (Schmitt et al., 2001 p.60). The VLT models a receptive form-meaning recognition format and measures five incremental vocabulary frequency bands: 2000, 3000, academic words, 5000, and 10000 frequency-level words. Each level has ten clusters, and each cluster has six words that need to be recognised with the right definitions. This helps to gauge learners' mastery of each of the five frequency levels, thereby providing an estimation of the overall vocabulary size. The validity and reliability of the VLT have been validated by Schmitt et al., Leontjev et al. (2001). Thus, the VLT is well placed to assess the Chinese university EFL learners' receptive size of word knowledge.

L2 WRITING TASKS

The L2 writing test selected two writing questions from the writing section of the International English Language Testing System (IELTS), task two in the academic module. One writing question was about the popularity of university education, and the other was about university online learning. Participants were required to write a 250-word essay for each of the two questions independently. In order to control the topical effects, which mean the differences in background knowledge that might affect the writing outcome, the two writing questions were selected with a similar topic of university learning so that participants may have similar background knowledge (Weigle, 2002).

DATA COLLECTION PROCEDURE

Ten students who were not among the participants attended the pilot study, ensuring that the target words and time allotment were feasible. In the main study, the six vocabulary tests were administered on two consecutive days to avoid test fatigue, and the testing sequence is shown in Table 4. The VLT is independent of the other depth tests and was the first one to be done. The depth tests were arranged from the hardest to the easiest to minimise the cross-test effects. The participants were not told that there would be another set of tests the next day to ensure the tests were low-stakes. They were prohibited from using dictionaries or other materials during the vocabulary tests. The researcher proctored the tests and collected each test before handing out the next one.

One week after the vocabulary tests, the participants were assigned the first writing task, in which they were required to naturally use at least five of the ten prompted target words in their writing. After a week's interval, participants were assigned to write the second essay and use at least five of the other ten target words. They were given more than one hour to complete each writing task.

TABLE 4. Vocabulary component tests sequence

Sequence	Tests	Time allotment (min)
Day 1	The VLT Test	30
	Collocation Test	20
	Derivative Test	25
Day 2	Association Test	20
	Form-recall Test	20
	Word-pair Test	20

THE SCORING SCALE

All receptive and productive word tests were scored dichotomously. No points were awarded for blanks, wrong spelling, non-existent words and words that did not match the syntactic contexts, yet grammatical errors such as wrong inflexions were ignored in the productive tests since participants may have partial knowledge of the word (Laufer & Nation, 1999). The reliability for all vocabulary depth tests is presented in Table 5. All Cronbach's alpha values were over .80, showing that all the depth instruments achieved high reliability.

TABLE 5. Reliability (Cronbach's alpha) for all vocabulary depth tests

Form recall	Word pair	Derivative	Association	Collocation
.830	.855	.842	.862	.860

The scoring of the target words was adapted from Zhong's (2016) 3-point scale. Three points were awarded for words used correctly and appropriately in L2 writing. Two points were given for words with minor mistakes, including one-letter spelling errors (such as adding or missing a letter, misspelling a letter, or mistakenly placing two adjacent letters) and grammatical errors (such as mistakes in word parts). One point was given for words that did not fit the context without interfering with the understanding of the intended meaning. No points were given if the word was misused in meaning and inappropriate in context.

TABLE 6. Total scores of all vocabulary knowledge tests

Tests	Total scores	Calculation
The VLT	150	3 items × 10 clusters at each level × 5 levels = 150
Form recall	20	20 items × 1 point each = 20
Word pair	20	20 items × 1 point each = 20
Association	57	20 items with max 57 keys = 57
Derivation	40	20 items × 2 points each = 40
Collocation	20	20 items × 1 point each = 20
Target words	60	20 items (10 items in each writing task) × max 3 points = 60

The present study employed the analytic scoring scale developed by Jacobs et al. (1981) to rate L2 writing. This rating scale is one of the most widely used analytical scales in EFL college-level writing (Weigle, 2002), and it may be clearer for L2 raters than holistic scoring. Five dimensions are included in this scoring scale, and the weights in each dimension were slightly modified in the present study as the research focuses on vocabulary and language use. The modified version included content (20%), organisation (15%), language use (30%), vocabulary (30%) and mechanics (5%).

Inter-rater reliability was verified for L2 writing and target word scores (see Table 7). The second rater was an experienced English instructor who has taught L2 writing for more than 15 years and was also working on her PhD program. She was trained by marking five scripts under the researcher's guidance before her independent rating. The scores awarded by the two raters for each essay and target words were compared to ensure reliability. The correlation coefficients of scores between the two raters ranged from .60 to .90, suggesting high consistency in scoring.

TABLE 7. Pearson correlation coefficients between the two raters

Writing task	Content	Organisation	Vocabulary	Language	Mechanics	Overall score	Target words
I	.867	.804	.850	.833	.658	.975	.965
II	.833	.789	.874	.776	.642	.936	.976

RESULTS

RESEARCH QUESTION 1: THE RELATIONSHIP BETWEEN THE SIX-WORD COMPONENTS AND L2 WRITING AND WORD USE

Descriptive statistics for scores on word knowledge components and L2 writing and word use are presented in Table 8. The percentages of these scores were also included for ease of comparison. Skewness and kurtosis of all tests were within the acceptable range of ± 3 , suggesting a normal data distribution. The collinearity diagnostics showed that all tolerance values were within the range of 0.3-0.6, and the VIF values were below 3, suggesting no collinearity issues among any pair of vocabulary tests.

Participants achieved the highest score on productive meaning with 87.9% of the total score (M = 17.58, SD = 2.666), followed by L2 writing at 77.8% (M = 155.59, SD = 8.429) and the VLT at 77.5% (M = 116.2, SD = 14.498). The lowest score was in collocation (M = 13.2, SD = 3.111) at 66%.

TABLE 8. Descriptive statistics for vocabulary components and L2 writing and word use ($n = 147$)

Test	Total	Mean (%)	Std. Deviation	Skewness	Kurtosis
The VLT	150	116.2 (77.5)	14.498	-0.156	0.295
Form recall	20	14.32 (71.6)	3.324	-0.646	0.499
Word pair	20	17.58 (87.9)	2.666	-1.365	1.299
Association	57	38.42 (67.4)	6.136	-0.599	-0.044
Derivation	40	30.01 (75.02)	5.380	-0.534	-0.293
Collocation	20	13.20 (66)	3.111	-0.234	-0.480
L2 writing	200	155.59 (77.8)	8.425	-.640	1.431
Target words	60	40.68 (67.8)	7.517	.084	-.580

The correlation matrix in Table 9 shows that all the correlation coefficients between word depth components are significant at $p < 0.01$ level ($r = .525 - .722$). Form recall and derivation showed the highest correlation among the individual depth components at $r = .722$. However, vocabulary size measured by the VLT was weakly correlated with the depth components (all below $r = .20$).

The correlations between all depth components and L2 writing and word use were close to or above large effect sizes. Derivation correlated most significantly with L2 writing ($r = .757$) and word use ($r = .768$), followed by form recall with L2 writing ($r = .617$) and word use ($r = .741$). The weakest relations were between the VLT and L2 writing ($r = .146$) and word use ($r = .153$).

TABLE 9. Correlation matrix between vocabulary components and L2 writing and word use

Test	1	2	3	4	5	6	7	8
1. VLT	1							
2. Form recall	.173**	1						
3. Word pair	.123**	.601**	1					
4. Association	.181**	.606**	.576**	1				
5. Derivation	.172**	.722**	.624**	.588**	1			
6. Collocation	.177**	.666**	.533**	.525**	.541**	1		
7. L2 writing	.146**	.617**	.609**	.501**	.757**	.491**	1	
8. Target words	.153**	.741**	.616**	.517**	.768**	.529**	.672**	1

$P^{**} < 0.01$; $p^* < 0.05$ (two-tailed); $n = 147$

RESEARCH QUESTION 2: CONTRIBUTIONS OF THE WORD KNOWLEDGE COMPONENTS TO WORD USE AND OVERALL L2 WRITING ABILITY

Hierarchical regression models were used to analyse the extent to which the word knowledge components contribute to L2 writing and word use, as shown in Tables 10 and 11. A boxplot was used to check for outliers, and nine cases were identified in the tests (two in VLT, four in word pair, two in form recall and one in association). All outliers below the Q1 cut-offs in VLT, word pair, form recall and association were considered to be mild. These data points were also included as they did not critically impact the overall data analysis.

The R^2 indicates the unique variance of L2 writing and word use that can be accounted for at each step by the corresponding predictor and the predictor(s) at the previous step (s). The R^2 change represents the addition of variance when a new predictor was entered. The local effect size in each model was calculated using Cohen's f^2 to express the specific effect brought by a given independent variable. According to Cohen (1988), the effect is small at $f^2 \geq 0.02$, medium at $f^2 \geq 0.15$, and large at $f^2 \geq 0.35$. The B values, standard errors (SE B) and Beta (β) extracted from Model 6, which includes all vocabulary components, are also reported here.

As shown in Model 1 (see Table 10), the first four steps were statistically significant at $p < 0.001$, with apparent improvements at each step. When form recall was entered, Step 2 explained 38.2% variance of L2 writing significantly at $p < 0.001$ with a large effect size ($f^2 = 0.921$). The addition of a word pair in Step 3 brought another 8.8% variance with a medium effect size ($f^2 = 0.224$). The R^2 change was significant with the entry of derivation at Step 4, increasing the variance explained to 60.7% with a large effect size ($f^2 = 0.394$). Form recall, word pair and derivation together accounted for more than half (58.6%) of the total variance. However, the addition of association and collocation cannot significantly increase the R^2 .

Overall, derivational knowledge had the largest contribution ($B = 0.179$, $\beta = 0.571$) to L2 writing performance, followed by productive meaning ($B = 0.063$, $\beta = 0.198$) and productive form ($B = 0.016$, $\beta = 0.023$). The VLT, association, and collocation were not significant predictors of L2 writing ability.

TABLE 10. Model 1 regression results of word components for L2 writing ($n = 147$)

Step & Predictors	B	SE B	Beta (β)	Sig.	R ²	R ² Change	effect size
Step 1				.000	.021	.021	0.054
VLT	.003	.024	.007	.078			
Step 2				.000	.382	.361	0.921
Form recall	.016	.023	.023	.000			
Step 3				.000	.470	.088	0.224
Word pair	.063	.023	.198	.000			
Step 4				.000	.607	.137	0.394
Derivation	.179	.026	.571	.000			
Step 5				.000	.607	.000	0.000
Association	.003	.028	.009	.960			
Step 6				.000	.608	.001	0.003
Collocation	.010	.020	.038	.608			

$p < 0.001$ (2-tailed); dependent variable = L2 writing

Similarly, the VLT alone in Model 2 (see Table 11) cannot significantly predict the target word use ($p = .065$). The addition of form recall at Step 2 significantly brought about the largest R^2 change, explaining 55% of the total variance at the $p < 0.001$ level, with the largest effect size at $f^2 = 1.622$. Word pair and derivation respectively added 4.5% at Step 3 and 7.8% at Step 4 to the overall variance explained. The R^2 changes were significant in these two Steps with medium effect sizes ($f^2 = 0.138$ and 0.240 , respectively). However, with association and collocation entered, neither the group models at the last two steps nor the newly added predictors were significant in predicting the target word use.

Taken together, the derivational forms contributed most to productive word use in writing ($B = 0.410$, $\beta = 0.441$). Form recall was the second strongest contributor ($B = 0.275$, $\beta = 0.365$), followed by productive meaning ($B = 0.145$, $\beta = 0.154$). The VLT, association and collocation cannot significantly predict the use of the target word.

TABLE 11. Model 2 regression results of word components for target word score ($n = 147$)

Step & Predictors	B	SE B	Beta (β)	Sig.	R ²	R ² Change	effect size
Step 1				.000	.023	.023	0.071
VLT	.007	.064	.005	.065			
Step 2				.000	.550	.527	1.622
Form recall	.275	.061	.365	.000			
Step 3				.045	.595	.045	0.138
Word pair	.145	.063	.154	.000			

Step 4				.022	.673	.078	0.240
Derivative	.410	.070	.441	.000			
Step 5				.176	.675	.001	0.003
Association	.055	.077	.047	.457			
Step 6				.179	.675	.000	0.000
Collocation	.009	.054	.011	.875			

$p < 0.001$; $p < 0.05$ (2-tailed); dependent variable = target word score

DISCUSSION

The present study adopted a multi-component approach to vocabulary knowledge and examined the interactions between receptive size and productive depth components and their relationships with L2 writing and productive word use. It also examined the extent to which each word knowledge component contributes to lexical use in writing and the overall L2 writing ability.

The correlation analysis revealed large correlations between the five depth components, which significantly correlated with L2 writing and word use. The result adds empirical evidence to the previous assumption that vocabulary knowledge is a network of interrelated lexical components. This is consistent with a number of studies (González-Fernández & Schmitt, 2020; Schmitt, 1998; Zhong, 2016) that explored the internal structure of word knowledge and demonstrated the interrelatedness between different word components. The various knowledge types of a word do not stand independent but strongly interact with one another.

The result shows that the strength of correlations between different word components might vary markedly (Schmitt, 1998). The receptive form-meaning links (vocabulary size), as assessed by the VLT, indicate small correlations with the other depth components and L2 writing and word use. This may contrast with the strong relationships (ranging from .760 to .895) between the VLT and depth components in González-Fernández and Schmitt (2020). Yet, all productive depth components in the present study revealed strong associations with L2 writing and word use. This might explain why the recognition knowledge in receptive size and recall ability in productive depth are distinct constructs. The awareness of this distinction can be critical to the conceptualisation of vocabulary knowledge (González-Fernández & Schmitt, 2020). Coupled with the high mean score ($M = 77\%$) in the VLT, our results are closely related to previous claims that many Chinese EFL learners simply “know little about a large number of words” (Schmitt, 2014, p.915) and have “lots of words in their mental lexicon but with poor organisation” (Milton & Fitzpatrick, p.150). In other words, they cannot translate their large vocabulary size to profound vocabulary depth in productive word use and overall L2 writing proficiency.

Consequently, receptive size by the VLT could not significantly predict target word use and L2 writing ability in the present study. This is surprising as it contradicts many other studies showing a strong predictive role of word size in writing performance (Dabbagh & Janebi Enayat, 2019; Stæhr, 2008; S. Wu et al., 2019). That said, it cannot be concluded that the receptive recognition knowledge makes no contribution to L2 writing because the R^2 change is small but still significant in the regression models. The test administration in the present study may be responsible for the discrepancy with previous studies. The VLT was administered online, which may give the participants chances to check answers to the form-meaning recognition items and obtain higher scores than expected.

Productive form and meaning knowledge elicited from form recall and word pair explain approximately half of the variance in L2 writing and word use. This is expected since productive form and meaning connections, referred to as active vocabulary by Laufer and Goldstein (2004),

are the most fundamental components in language production. The most striking difference between native and non-native speakers is the number of words they can control in free language production, and a large active vocabulary also contributes to a higher level of lexical richness in writing (Crossley et al., 2015). The form and meaning links in productive contexts can be operationalised as form recall ability by the PVLTL (Schmitt, 2014) and L1-L2 word mapping knowledge (Nation, 2022). This has been demonstrated by the large portion of unique variances accounted for by the two components in L2 writing and word use in the present study.

Compared to form recall and word pair, derivation, association, and collocation reveal considerably fewer predictions in L2 writing and word use. Yet, the addition of 13.8% R^2 change brought by the three depth components in L2 writing is statistically significant. Moreover, the six-component model, including the VLT, can significantly predict L2 writing with a large global effect size $f^2= 1.596$, meaning that the variance explained by all vocabulary components together is roughly 1.6 times higher than the unexplained variance. This result, to a certain extent, lends empirical support to the theoretical frameworks by Nation (2013, 2022) and Coxhead (2007), which theorise that vocabulary knowledge includes various components at receptive and productive levels and that all the components, beyond the basic form and meaning links, are worth investigating in L2 writing performance.

The productive derivative production has the highest correlations with L2 writing and word use (both above .70), and the R^2 changes it brings are also large and significant. This is consistent with previous findings that even if L2 learners have uneven mastery of productive derivative knowledge, this knowledge accounts for more than 10% of the lexical errors identified in L2 university writing (Schmitt & Zimmerman, 2002). The correlations in the present study are stronger than those of Leontjev et al. (2016) findings, which show moderate to large relations between derivation and writing proficiency. Interestingly, judging from the Beta (β) values in the present study, derivation contributes most to both L2 writing and word use. The values are even higher than productive form knowledge, even if form recall explains the most R^2 with the largest effect size. This may be due to the shared variance between derivation and form, as morphological knowledge is regarded as part of form knowledge (Nation, 2022). This implies that knowing more derived forms of the word is as, if not more, important as form recall for productive use. Derivative knowledge can be a reflection of form and needs to be activated to determine the word position and sentence structure in written production (Nation, 2022).

However, the results show that productive association and collocation contribute little to L2 writing and cannot significantly predict L2 writing and word use individually. This can partly be explained by the large portion of shared variance. Productive meaning shared more than 50% of variance explained by association in L2 writing and word use, and the majority of variance explained by collocation was shared by productive form. When they entered the models, a large share of R^2 was already accounted for by productive form and meaning. The non-significant association may also be attributed to the writing output, which may not reflect the process of recalling and comparing different synonyms to select the most appropriate words in L2 writing. As for collocation, our results are consistent with previous findings showing that even advanced learners lack the capacity to use collocations in their writing (Laufer & Waldman, 2011). Laufer and Waldman (2011) explained that this difficulty is mainly caused by semantic transparency and L1 transfer of collocations. This is especially so in the Chinese context, as the Chinese-L1 and the word-for-word meaning translation have a positive role in processing and writing English collocations (Y. P. Wu et al., 2024). These explanations also hold for this study. Nevertheless, the

significant R^2 changes and group models in L2 writing brought by association and collocation indicate that the two-word depth components should not be overlooked in L2 writing proficiency.

CONCLUSION

The present study employed a multi-component method to investigate the role of multiple size and depth word components in L2 writing and word use. The study empirically confirms the close interrelations between multiple-word knowledge components, as illustrated in the theoretical frameworks of Nation (2013, 2022) and Coxhead (2007). The predictive results demonstrated that while productive form and meaning aspects are still critical for L2 writing, productive morphological knowledge plays the most pivotal role in word use and overall writing ability. This serves as a pedagogical reminder that language teachers need to give priority to explicit derivational instructions when teaching a word, especially for productive goals. Nevertheless, association and collocation cannot significantly predict productive skills, but they still have large correlations with them and significantly improve the regression models in L2 writing. The receptive size, as measured by the VLT, has no relationship with L2 writing and lexical proficiency. This suggests that receptive recognition of form and meaning knowledge is insufficient for a writing context, even with a large vocabulary size. The present study advocates that Chinese EFL word teaching should not focus on receptive form-meaning links but rather dig deeper into the productive use of multiple word components, including derivations, synonyms and collocations.

LIMITATIONS AND IMPLICATIONS

One of the limitations of the present study is the mutual interference between the word knowledge components due to their close relationships, though the tests were arranged in a considered order to avoid providing hints for test-takers. The administration of some tests online may also offer chances for test-takers to find the answers from different sources, which might inflate their scores. In addition, the L2 writing tasks that were devised to assess all the word components measured may not be able to activate and reflect all of them. For example, association was not revealed in the writing output and form recall was not needed in the controlled writing tasks since the target words were provided as prompt words. Future studies may design better productive tasks to measure these two-word components, such as free writing tasks for form recall. Finally, this is a cross-sectional study to capture word knowledge and productive skills within a certain time point, and it is unable to observe the dynamic changes in vocabulary knowledge and L2 writing proficiency. Thus, longitudinal studies in terms of the relationship between vocabulary knowledge and L2 writing are recommended in future research.

ACKNOWLEDGEMENTS

This research project was financially supported by Mahasarakham University.

REFERENCES

- Bestgen, Y. (2017). Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence. *System*, 69(2017), 65-78. <https://doi.org/10.1016/j.system.2017.08.004>
- Choo, L. B., Lin, D. T. A., Singh, M. K. M., & Ganapathy, M. (2017). The Significance of the Academic Word List among ESL Tertiary Students in a Malaysian Public University. *3L: The Southeast Asian Journal of English Language Studies*, 23(4), 56-65. <http://doi.org/10.17576/3L-2017-2304-05>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum.
- Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly*, 34(2), 213-38. <https://doi.org/10.2307/3587951>
- Coxhead, A. (2007). Factors and aspects of knowledge affecting L2 word use in writing. In P. Davidson, C. Coombe, D. Lloyd, & D. Palfreyman (Eds.), *Teaching and Learning Vocabulary in Another Language* (pp. 331-342). Dubai: TESOL Arabia
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2015). Assessing Lexical Proficiency Using Analytic Ratings: A Case for Collocation Accuracy. *Applied Linguistics*, 36(5), 570-590. <https://doi.org/10.1093/applin/amt056>
- Dabbagh, A., & Janabi 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>
- González-Fernández, B., & Schmitt, N. (2020). Word Knowledge: Exploring the Relationships and Order of Acquisition of Vocabulary Knowledge Components. *Applied Linguistics*, 41(4), 48-505. <https://doi.org/10.1093/applin/amy057>
- Jacobs, H., Zingraf, A., Warmuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Kim, M., Crossley, S. A., & Kim, B. K. (2022). Second language reading and writing in relation to first language, vocabulary knowledge, and learning backgrounds. *International Journal of Bilingual Education and Bilingualism*, 25(6), 1992-2005. <https://doi.org/10.1080/13670050.2020.1838434>
- 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, P. (1999). A vocabulary-size test of controlled productive vocabulary. *Language Testing*, 16(1), 33-51. <https://doi.org/10.1177/026553229901600103>
- Laufer, B., & Waldman, T. (2011). Verb-Noun Collocations in Second Language Writing: A Corpus Analysis of Learner's English. *Language Learning*, 61(2), 647-672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Leontjev, D., Huhta, A., & Mantyla, K. (2016). Word derivational knowledge and writing proficiency: How do they link? *System*, 59(2016), 73-89. <https://doi.org/10.1016/j.system.2016.03.013>
- Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, 36(5), 611-34. <https://doi.org/10.1093/applin/amu007>
- McLean, S., Stewart, J., & Olaf Batty, A. (2020). Predicting L2 reading proficiency with modalities of vocabulary knowledge: A bootstrapping approach. *Language Testing*, 37(3), 389-411. <https://doi.org/10.1177/0265532219898380>
- Milton, J., & Fitzpatrick, T. (Eds.). (2014). *Dimensions of vocabulary knowledge*. Palgrave Macmillan.
- Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.)*. Cambridge University Press.
- Nation, I. S. P. (2022). *Learning vocabulary in another language (3rd ed.)*. Cambridge University Press.
- 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. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing* (pp. 209-227). John Benjamins.
- Read, J., & Dang, T. N. Y. (2022). Measuring depth of academic vocabulary knowledge. *Language Teaching Research*, 1-27. <https://doi.org/10.1177/13621688221105913>
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281-317. <https://doi.org/10.1111/1467-9922.00042>
- 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. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching*, 52(2), 261-274. <https://doi.org/10.1017/S0261444819000053>

- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88. <https://doi.org/10.1177/026553220101800103>
- Schmitt, N., & Zimmerman, C. B. (2002). Derivative Word Forms: What Do Learners Know? *TESOL Quarterly*, 36(2), 145-171. <https://doi.org/10.2307/3588328>
- 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>
- 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. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/263457>
- Waluyo, B., & Bakoko, R. (2021). Academic Vocabulary Used by High School Students in Essays and Its Relation to English Proficiency. *3L: Southeast Asian Journal of English Language Studies*, 27(4).
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge University Press.
- Wu, S., Dixon, Q., Sun, H. T., & Zhang, P. (2019). Breadth or depth: the role of vocabulary in Chinese English Language beginning writers' development. *International Journal of Bilingual Education and Bilingualism*, 24(9), 1356-1372. <https://doi.org/10.1080/13670050.2019.1572066>
- Wu, Y. P., Halim, P. A., & Hamzah, M. H. (2024). Assessing the Roles of L1 and Meaning Knowledge in Processing and Production of English Collocations among Chinese EFL Learners. *3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies*, 30(1), (2024). <http://doi.org/10.17576/3L-2024-3001-13>
- Yanagisawa, A., & Webb, S. (2020). Measuring depth of vocabulary knowledge. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 371-386). Routledge.
- Zhong, H. (2016). The relationship between receptive and productive vocabulary knowledge: A perspective from vocabulary use in sentence writing, *The Language Learning Journal*, 46(4), 357-370. <https://doi.org/10.1080/09571736.2015.1127403>

APPENDIX: EXAMPLES OF TEST INSTRUMENTS

FORM RECALL

Instructions: Complete the underlined words in the sentence. An example has been done for you.
The hotel wants to expand its business by adding two more stores.

WORD PAIR

Instructions: Complete the word according to the Chinese translation. An example has been done for you.

扩张/增长 --- e_____

PRODUCTIVE DERIVATIVE

Instructions: Fill in the blanks with the appropriate forms of the prompt words. Please note the context of the sentence.

determine

Noun: He fought the illness with courage and _____.

Adjective: He was _____ that the same mistakes would not be repeated.

ASSOCIATION

Instructions: Tick off words in the below box with the similar meanings as the underlined word in the sentence. There may be 2-3 keys in the box. DO NOT choose more than three words.
The water froze inside the pipe, causing it to expand and burst.

enlarge	swell	contract	dilate	explode
---------	-------	----------	--------	---------

PRODUCTIVE COLLOCATION

Instructions: Complete the sentences with an appropriate collocation. Collocation means phrases in which the word given always appears with other words in sentences. The first letters of the words (including prepositions) you have to use to make the collocation have been provided. These target collocations you have to complete include different types:

noun+preposition, adjective+noun, and verb+noun, among other combinations. The Chinese sentence has prompted the target collocation.

读写结合模式可以说是一种新的实践性很强的语言教学途径。

Many teachers believe that this can be a new, practical a_____ t_____ language teaching.

L2 WRITING TASK

Instructions: Please write an argumentative essay on the following question. You are required to use at least **five** words (the more, the better) randomly selected from the below box. Please

integrate the words you choose in your essay in the most natural manner, cohesively and grammatically. You can use any derived form of the word (i.e., verbs to adjectives) you choose and mark them in your essay with a circle or underline. You should write 250-300 words within one hour (60 minutes).

Writing task one: Competition for places at university is increasing. Why do more and more people want to study at university? Is this a positive or negative development?

indicate	achieve	justify	prior	exposed
intelligence	rely	consume	devoted	lecture

Writing task two: Today, millions of university students have to enroll in online learning for higher education. Colleges and universities offer e-learning programs and courses. Do you agree or disagree with the popularity of online learning?

approach	persist	detect	access	expand
stimulate	voluntary	schedule	restrain	inclined