

Assessing the Role of Vocabulary Knowledge in Developing EFL Learners' Writing Skills: Implications for Intentional and Incidental Vocabulary Learning

Linda H. F. LIN

The Hong Kong Polytechnic University, HKSAR

Vocabulary knowledge is essential for language learning. This is particularly so for English as a second language (L2) and foreign language (EFL) learners, whose affordances of input and output opportunities may be limited. A rich contextual learning environment is instrumental for these language learners because it enables incidental learning to take place. However, little research has investigated the role of vocabulary knowledge in developing EFL learners' writing proficiency. The current study, drawing upon two vocabulary tests, one writing test, three focus group interviews, and two case studies, addresses this gap from multifaceted perspectives. The findings of the study provided empirical evidence of the critical role of vocabulary knowledge in EFL learners' writing proficiency. The results also revealed the centrality of a contextual learning environment in developing EFL learners' writing skills. The paucity of such an environment limited the learners' writing experience, hampered their confidence, and caused their misconceptions of what counts as good writing. This study contributes to research on the relationships between learners' vocabulary knowledge and writing proficiency. It has significant implications for intentional and incidental vocabulary learning and profound implications for EFL teaching pedagogy.

Introduction

This study investigates the role of vocabulary learning in developing EFL learners' writing skills. It examines this issue by taking into consideration contextual factors. Contextual factors are closely related to learners' in-classroom (e.g., adopted teaching approaches and classroom ethos) and out-of-class learning environments (e.g., high-stakes examinations and the richness of input and output opportunities to use the target language). These factors could "impinge on the dynamics of language teaching and learning in various ways, for example, by creating or withholding opportunities to use and experience the target language and by shaping learner perceptions, learning strategies, and classroom behaviors" (Hu, 2003, p. 303). The paucity of conducive contextual factors in China has led to many learners' dependence on rote learning, which focuses mainly on memorizing vocabulary and grammar rules (Jin & Cortazzi, 2011). This learning style is manifested in lexical learning, where many learners focus on decontextualized word lists, the most common approach used in intentional vocabulary learning (Webb, 2020). Although this approach plays an important role in enhancing students' receptive skills in comprehending written and aural texts, its value in improving learners' productive skills in producing written and oral texts is rather limited. This is because learners' productive mastery contributes most to their ability to appropriate lexical production (Schmitt, 2014). Productive mastery of words requires incidental learning, which depends on frequent encounters with words in natural language contexts, that is, through reading and listening (Webb, 2020).

Compared to speaking, writing places a higher demand on learners' lexical knowledge due to the complexity of different genre conventions. This higher demand requires more incidental learning than intentional learning. Words acquired through incidental learning can be more easily activated and more appropriately used in different genres. A singular lack of incidental learning can severely impact learners' writing proficiency. This impact is particularly evident in EFL learners who are "linguistic outsiders" of the target language (Folse, 2004, p. 1). However, there is a dearth of research on the role of vocabulary knowledge in developing EFL learners' writing proficiency. This study intends to bridge this gap by examining how the quality of the vocabulary knowledge of these learners affects their writing skills. The results of the study provide valuable insights into incidental and intentional vocabulary

learning and thus have significant implications for teaching and learning vocabulary for EFL learners.

Literature Review

Vocabulary Knowledge in Language Learning

The centrality of vocabulary knowledge has been well documented in the related literature. “Language learning is largely lexical learning,” according to Gass and Selinker (2008, p. 173). Lexical knowledge is “fundamental to all language use,” as asserted by Schmitt et al. (2017, p. 213). Qian and Lin (2020) conceptualized this strong connection by contending that vocabulary knowledge is a key prerequisite for successful language learning. This prominent role of vocabulary knowledge is even more evident in EFL and L2 learners due to contextual factors. The lexical development of first language (L1) learners is highly contextualized given the affordances of their extensive and natural exposure to the target language. The lexical learning of L2 and EFL learners, on the other hand, is complex and thus more onerous because of their limited natural exposure to the target language. This complexity is fully manifested in the process of lexical development, which can be seen as a continuum where learners start with superficial familiarity with a word and end with an ability to use the word appropriately in free production (Færch et al., 1984). Progression through this continuum is the development of the quality or depth of lexical knowledge. Vocabulary depth plays a more important role than vocabulary size (the quantity of word knowledge) in developing learners’ productive skills (Schmitt et al., 2017).

The size dimension of a learner’s lexical knowledge is single-faceted and thus quantifiable and can be tested with size tests, such as Nation’s (1990) Vocabulary Levels Test (VLT), which is “the *de facto* standard vocabulary size test” (Meara & Alcoy, 2010). The depth construct of a learner’s lexical knowledge, on the other hand, is multifaceted and covers a range of word knowledge in one’s lexicon, resulting in challenges in assessing this construct of word knowledge. To counter this difficulty, most researchers focus on key aspects, such as grammatical functions, collocations, and constraints on use. This can be seen in the Word Associates Test (WAT; Read, 1998).

Apart from testing lexical knowledge, both the VLT and the WAT have been used to predict EFL and L2 learners’ language abilities. Stæhr

(2008), for example, compares 88 EFL learners' examination grades with their vocabulary size and identifies a strong correlation between the two, particularly in the grades for reading ($r = .83, p < .01$) and writing ($r = .73, p < .01$). This study confirms the result of a fundamental study by Qian (1999), which finds that vocabulary size has a high and positive correlation with L2 reading comprehension abilities. Meanwhile, Qian provides empirical evidence showing that depth of vocabulary knowledge makes a unique contribution (11%) to predicting L2 learners' ability in reading comprehension, over and above the contribution already made by the size of the vocabulary. Another empirical study in the same vein is Stæhr (2009). This investigation focuses on the relationship between learners' vocabulary knowledge (both in breadth and depth) and their listening comprehension and documents a strong correlation between the two. Meanwhile, the study finds that depth of vocabulary knowledge adds only 2% to the variance already explained by vocabulary size in predicting listening comprehension (p. 592). While it is possible to explain this limited change with the "fleeting nature" of spoken discourse (Qian & Lin, 2020, p. 70), which leads to high pressure on learners' short-term working memory, it is worth noting that both aforementioned studies focus on receptive skills. Since vocabulary depth defines the quality of learners' word knowledge or how well they can activate words for appropriate language production, it is highly possible that the predictive power of vocabulary depth on productive skills (e.g., writing) is higher than that of vocabulary size. This is a research area that is warranted in lexical research (Yanagisawa & Webb, 2020).

Writing skills depend largely on three key constructs: lexis, grammar, and cohesion (Crossley, 2020). Of the three constructs, the first two, word use and syntactical structures, are inextricably intertwined. This relationship is most manifest in the appearance of lexicogrammar (Halliday & Matthiessen, 2013), a linguistic approach focusing on the integration between vocabulary and grammar in language production. This explains why the first two key linguistic elements in judging writing quality are "*words* writers produce and the structure those *words* are placed" (Crossley, 2020, p. 417). However, an important factor to consider in studies on the association between lexical knowledge and writing skills is learners' approaches to word learning, that is, incidental, intentional, or an effective combination of both. This consideration is based on the assertion that learners' success in activating appropriate and

contextual words in writing depends on the affordances of their input and output opportunities (Laufer & Paribakht, 1998).

Intentional and Incidental Vocabulary Learning

The vocabulary learning approaches taken by L2 and EFL learners vary according to contextual factors of language learning. There are two main vocabulary learning approaches, incidental and intentional. The former rests on repeated encounters with a word in natural language contexts, that is, through reading and listening, and thus allows the word to be learned gradually over time (Webb, 2020). This is the major lexical acquisition mode for L1 learners. The latter depends largely on connecting the L2 form of a word to its L1 meaning and thus often takes place in the form of memorizing words with flashcards and/or decontextualized word lists.

Notwithstanding the various advantages of intentional learning, such as facilitating the learning of formulaic expressions, allowing fast gains in technical vocabulary, and more importantly, laying groundwork for incidental learning to take place (Liu & Nation, 1985), incidental learning is essential for L2 and EFL learners. This is because incidental learning develops links and strengthens the interconnectedness of words in one's mental lexicon (Pavlenko, 2009; Singleton, 2007). Learners normally gain the most frequent meaning of a word first and then encounter the secondary and peripheral meanings of the word gradually over time. Repeated word encounters in reading and listening allow for various aspects of lexical development (Webb, 2020) and lead to learners' automaticity and appropriateness in language use, the ultimate lexical development goal for most language learners (Gu, 2020). This is possibly why in Nation's (2007) four strands of lexical development framework, three of them (meaning-focused input, meaning-focused output, and fluency) need to be undertaken through incidental learning, and only one, language-focused learning, can be conducted via intentional learning.

Contextual Factors in China

Notwithstanding the centrality of incidental vocabulary learning in language development, many EFL learners still focus on intentional learning. This is because learning approaches are largely affected by the

contextual factors of language learning. If the contextual factors are uncondusive, many learners have little choice but to resort to intentional learning. This is the case in China (Lin, 2019).

The contextual factors in China are uncondusive (Yu et al., 2018). English is a foreign language in the country. There are thus very limited input opportunities for English learners, for example, English-medium television, newspapers, and social media sites. A large majority of learners rely on vocabulary lists, textbooks, and even examination-preparation books for English learning (Lin, 2019; Zhao, 2016). Their output opportunities are hardly better. Most learners rarely have opportunities to be engaged in meaningful English conversations (Liu & Jackson, 2011). In other words, English is rarely used in the community (Butler, 2014). With the learning context in China as such, learning English mostly becomes instrumentally motivated, that is, to pass various high-stakes English examinations, such as the National Matriculation English Test (NMET), the English test in the national university entrance test battery. To achieve a high score on the NMET, many teachers provide intensive vocabulary dictations and grammar drills in class, and students respond by memorizing a large number of words from decontextualized word lists (Qian & Lin, 2020). Such intensive training largely reinforces many learners' beliefs that learning English is a process of accumulating vocabulary and understanding grammar rules. The consequence is that even learners who are aware of the importance of contextual vocabulary use also focus mainly on the meaning retention of words via intentional learning. This lack of incidental learning could severely affect the development of vocabulary in learners' mental lexicon, which in turn impacts their writing skills.

Research Aim and Research Questions

Given the aforementioned discussion on the need for research into the association between learners' vocabulary knowledge and writing skills and the required attention to learners' approaches in such studies, this study investigates the role of vocabulary knowledge in developing the writing skills of EFL learners in China. Drawing upon the framework of Qian (1999) and Stæhr (2009), it aims to answer three research questions (RQs):

- RQ1. How does learners' vocabulary knowledge correlate with their writing proficiency?
- RQ2. Compared with vocabulary size, is the depth of vocabulary knowledge a more reliable predictor of learners' writing proficiency?
- RQ3. To what extent does the lack of incidental vocabulary learning affect learners' writing experience and writing proficiency?

Methodology

Research Design

Both quantitative and qualitative data were gathered for this study. The main instruments for collecting quantitative data involved two vocabulary tests and a writing test. The qualitative data was gathered through three focus group interviews and two case studies. The information elicited by these methods provides multiple and yet complementary perspectives on the impact of the paucity of vocabulary knowledge on the writing skills of learners in China.

Participants

A total of 83 first-year undergraduates in a major university in Hong Kong were invited to participate in the study. All participants received their primary and secondary school education in Mainland China, and most of them were elite secondary school graduates. Since English is the medium of instruction in Hong Kong universities, a requirement for English proficiency is imposed on students. For candidates from Mainland China, a minimum of an overall mark of 120 (out of 150) on the NMET is needed. This means that only approximately the top 5% of secondary school graduates in the Mainland may have the opportunity to enter a Hong Kong tertiary institute. Detailed profiles of the participants can be seen in Table 1.

Table 1. Profiles of Participants

Number of participants	Gender ratio	Age range
83	35 males 48 females	16–22

Vocabulary Tests

The two vocabulary tests were the VLT and the WAT. The VLT gauges learners' receptive vocabulary by assessing their performance in word-definition matching, as in the example below:

Sample Question 1

1 adequate	
2 internal	_____ 1 _____ enough
3 mature	_____ 3 _____ fully grown
4 profound	_____ 5 _____ alone away from other things
5 solitary	
6 tragic	

The WAT is the “most-used measure of depth” of learners' lexical knowledge (Schmitt, 2014, p. 938). Each item in the test contains a target word, which is an adjective, and eight other words. Of the eight words, four are either semantically or collocationally associated with the tested word, and the other four are not related to the stimulus word in any sense (see the example below).

The VLT at the 3,000 and 5,000 frequency levels was employed. The rationale for only adopting these two levels is that the VLT on the academic word list covers vocabulary from the second to the fifth frequency levels and is thus not targeted at words at a particular level. The 2,000 level is too basic for EFL students at the university level, while the 10,000 level is far too difficult for most of them. Several related studies also lend support to this choice of tests. Schmitt et al. (2004), for example, deployed the VLT only at the 3,000 and 5,000 frequency levels to investigate the formulaic language used by EFL students at the University of Nottingham because “the 2,000 level was deemed too basic for the relatively advanced English for Academic Purposes (EAP) students, while the 10,000 level was still considered quite difficult” (p. 59). Laufer and Ravenhorst-Kalovski (2010) also excluded the 10,000 level when measuring the lexical threshold for reading comprehension of students in a university in Israel because this level of the test was “considered far too difficult in view of the background they had in English” (p. 21).

Sample Question 2

dense	crowded hot noisy thick
	forest handle smoke weather

A number of test sessions were arranged so that the participants could attend the tests at their own convenience. During each session, all participants were asked to complete the two vocabulary tests and write an essay of approximately 500 words. All these procedures were administered under controlled conditions, that is, in a classroom environment monitored by an instructor. No dictionaries or electronic devices were allowed in the process. All participants were given the same amount of time for each of the test sessions.

Essay Writing

Aligned with the Prompts Design Guidelines by Hamp-Lyons and Kroll (1996), the prompts for essay writing in the current study were “as brief as clarity allows” (p. 60). The respondents were able to choose from a pool of five topics. This was to maximize the opportunity for each participant to write on a topic she or he was familiar with and comfortable writing about. One example is as follows:

Discuss causes of the wealth gap in China and suggest ways to address this issue.

The essay scoring for the present study consisted of two stages. Holistic scoring was adopted in the first stage. This method “rates the paper as a whole (holistically) and assigns the paper a single score within a given range on scales” (Reid, 1993, p. 291). This evaluation method is often used for placement tests, whose major objective is to separate learners into different levels according to their writing proficiency. Since the main objective of the writing test for the current study is similar, this assessment method was adopted. To enhance the interrater reliability of the rating, two measures were undertaken. First, two experienced language instructors rated each essay, and both raters evaluated the essays according to the writing band descriptors for IELTS (Task 2), which have undergone careful research and piloting for their reliability and predictive validity (Shaw & Falvey, 2008). Second, after all the scripts were rated, an initial round of intraclass coefficient (ICC) analysis, which measures the consistency between the raters’ judgments, was performed. The scores awarded by the two raters for each essay were then compared. If they were the same or different on one scale, they were accepted. If the scores were diversified by two or more scales, they were graded for a second time by both raters; this time, using analytic scoring, a method in which different aspects of a script are evaluated independently and each given a

score. This detailed scoring procedure requires raters to “attend to the multidimensionality” of a script and thus allows them to make “more valid judgments” about the writing (Hamp-Lyons & Kroll, 1996, p. 62).

After the analytic scoring process, scripts with markedly different scores by the two raters were discussed until the two raters reached a consensus on the final score. After this process, a second round of ICC analysis was performed to examine the improvement in interrater consistency.

Focus Group Interviews

While the quantitative data derived from the above vocabulary tests and essay writing provides answers to RQ1 and RQ2, the answer to RQ3 requires qualitative data gathered from focus group interviews and case studies. There were a total of three focus group interviews for the study. All 83 participants were invited to the interviews, but only 21 of them accepted the invitation. The 21 voluntary participants were then divided into three groups and interviewed at different times: at the beginning, in the middle, and at the end of the first semester of their university studies. This study design had three major advantages. First, it ensured that each interview would not last very long, and that all the participants at each session had sufficient time to thoroughly express their views. Second, it allowed the researcher time to observe the participants’ language use and then ask further questions at the next interview. Third, the time span between the three interviews allowed the participants to see the differences between English learning in their secondary school, where intentional learning dominated, and in the university, where incidental learning was needed.

All three interviews lasted approximately one hour and were conducted in English. Each participant was invited to only one of the interviews. To elicit information from different perspectives, each interview focused on a specific theme (see Table 2).

Table 2. Themes, Participants, and Timing of Focus Group Interviews

Interview	Themes	Gender of participants	Timing
1	English learning experience in Mainland China—vocabulary	4 males 3 females	Beginning of semester
2	English learning experience in Mainland China—writing	4 males 2 females	Mid-semester
3	Difference in vocabulary learning and writing in secondary school and university	4 males 4 females	End of semester

Case Studies

While the analyses of the two vocabulary tests, essay writing scores, and focus group interviews provided a broad-brush landscape of the role of vocabulary knowledge in developing learners' writing skills, the case studies offered a portrait in miniature of such impacts. Two cases were carefully followed, with the first one, Jacky, from Duyun, a small city, and the second one, Mary, from Beijing, one of the most developed metropolitan cities in China.

Data Analysis

The analysis of the quantitative data included a Pearson's correlation test and a multiple regression test. The correlation test was performed to measure the correlation between the participants' vocabulary size (VS) and depth of vocabulary knowledge (DVK) and the results of their essay writing (EW). A multiple regression test was undertaken to verify the test result from the correlation test, specifically to identify if the depth of the learners' knowledge makes a unique contribution to their writing proficiency.

The analysis of the qualitative data mainly involved transcribing all three interviews for detailed coding and then theme identification. The coding followed Braun and Clarke's (2006) recursive process. It started with reading and annotating the transcriptions, which was then followed by identifying recurring themes and subthemes. NVivo, a specialized software for thematic data analysis, was also deployed to complement the above coding steps. This software helped to create visual connections between the recurring themes and subthemes and was particularly effective in coding and recoding the two main themes for the current study, vocabulary learning and writing skill development. Saldaña's (2021) framework on values coding, which views the identified themes through the lenses of values, attitudes, and beliefs of the respondents, was instrumental in the theme identification process.

Findings and Discussion

Association Between VS, DVK, and Learners' Writing Skills

Table 3 reveals the results of the two vocabulary tests. The mean score of VS was 42 ($\alpha = .85$) and that of DVK was 112 ($\alpha = .88$). This

indicates that the participants achieved, on average, 70% of the correct answers on both tests. These scores indicate that the learners possessed an acceptable level of vocabulary knowledge.

Table 3. Scores on VS and DVK (N = 83)

Test	Maximum possible score	Mean	SD	Score range	Cronbach's alpha reliability
VS	60	42	7.36	19–58	.85
DVK	160	112	14.16	57–136	.88

Note. DVK = depth of vocabulary knowledge; VS = vocabulary size.

However, their EW scores cast doubt on this conclusion. The mean score of the 83 respondents was only 4.11 ($SD = 1.32$), less than half of the maximum score of the writing test, which was 9. As lexical use is the key determinant of writing quality (Crossley, 2020; Gass & Selinker, 2008), especially for texts written by L2 learners (Llach, 2011; McCarthy, 1990), the scores on the two vocabulary tests, particularly those on DVK, may not represent the lexical skills in their language production. To test this possibility, two types of analyses were conducted. The first was a Pearson product-moment correlation test to identify associations among scores on VS, DVK, and EW. Table 4 shows that the VS scores were closely associated with EW scores ($r = .68, p < .01$), and the correlation level was similar to that of VS with reading comprehension ($r = .78$; Qian, 1999) and listening comprehension ($r = .70$; Stæhr, 2009). However, the DVK scores were only weakly related to the writing scores ($r = .28, p < .05$). This result is rather unexpected, as it was much lower than the correlations of DVK with reading comprehension ($r = .82$; Qian, 1999) and listening comprehension ($r = .80$; Stæhr, 2009). What was even more surprising was that the correlation between VS and DVK was only $.34 (p < .01)$. Although the correlation at this level is significant, it is much lower than the results in Qian (1999) and Stæhr (2009), at $.82$ and $.80$, respectively. These results suggest that many of these learners might have obtained effective test-taking strategies and thus performed reasonably well on DVK. However, these scores did not represent the lexical skills in their writing production.

Table 4. Pearson Correlations Among VS, DVK, and EW (N = 83)

Test	EW	DVK
VS	.68**	.34**
DVK	.28*	

Note. DVK = depth of vocabulary knowledge; EW = essay writing; VS = vocabulary size.

*Significant at .05. **Significant at .01.

To further verify this result, the current study undertook a second analysis, a regression test. To perform this analysis, VS was first manually entered into the regression equation. This step showed that VS alone explained 11% of the variance in these learners' writing efficiency ($p < .01$). When DVK was entered into the model later, the R^2 changed to .12, suggesting that the depth dimension of the learners' lexical knowledge added 1% to the variance already accounted for by the learners' VS. This 1% change added only a very limited increase to the variance already explained by VS and is statistically insignificant (Table 5).

Table 5. Regression Results with VS and DVK as Independent Variables (N = 83)

Step	Procedure	Variable	Status	R^2	R^2 change
1	Forced entry	VS	In	.11*	
2	Forced entry	DVK	In	.12*	.01

Note. DVK = depth of vocabulary knowledge; VS = vocabulary size.

*Significant at .05.

The above results provide answers to RQ1 and RQ2 of the current study. First, the learners' VS scores significantly correlated with their writing scores, a finding closely in line with both Qian (1999) and Stæhr (2009). However, their DVK scores were only weakly associated with their writing performances, a result rather different from Qian (1999) and Stæhr (2009) since both studies revealed a strong correlation between DVK and L2 learners' language skills. Second, the predictive power of vocabulary depth on learners' writing skills was only 1%, much lower than that of VS (11%). This statistically insignificant 1% change that DVK added to the variance already accounted for by VS was even more surprising. It was lower than the 11% added to reading comprehension reported in Qian (1999) and the 2% added to listening comprehension reported in Stæhr (2009). More importantly, the changes in both studies are statistically significant.

The above analysis and the weak correlation between the learners' VS and DVK scores, which are much lower than the corrections found in Qian (1999) and Stæhr (2009), indicate that the learners' performance on DVK had little association with their writing skills. In other words, their lexical knowledge remained largely at the receptive level. Some learners might have gained the depth of lexical knowledge of a number of words, but their knowledge has not been translated into the ability to use lexical items fluently. This situation suggests that these students are learners who "know *little* about a *large* number of words" (Schmitt, 2014, p. 915). In Milton's explanation, they belong to those "with lots of words" in the mental lexicon, but the words are in "poor organization" (2009, p. 150). One possible explanation for this poor organization is the lexical learning approach adopted by most of the learners, which may have severely impacted the interconnectedness between words in their lexicon. This explanation is drawn upon findings in prior studies that learners' ability to activate appropriate words in writing is largely determined by their incidental lexical learning (Laufer & Paribakht, 1998; Schmitt, 2014). To verify this possibility, the study conducted three focus group interviews.

Focus Group Interviews

The three focus group interviews were conducted at different times of the semester. Each interview, although with a different theme, was focused on their experience in lexical learning and writing, which formed the two main themes of the coding process. A number of subthemes also appeared under the two main themes. The analysis of the themes and subthemes showed that intentional learning was by and large the dominant approach to English learning before their arrival in Hong Kong. This finding confirmed the results of prior studies (Hu, 2003; Jin & Cortazzi, 2011; Lin & Morrison, 2021). Their adopted approach was determined by the affordances of input and output opportunities, as evidenced by their language training in school. The focus of the training was on skills needed for the NMET. The description of a respondent from a small city in northeast China during the second interview clearly illustrated the situation:

My classmates and I went to school seven days a week. We started at 7 am and finished... normally by 9 pm every day. That was... all through our six

years in the secondary school. We had English classes in the morning and afternoon. The teacher usually taught new content [a new text]... in the afternoon. He first of all go [went] through the vocabulary list, then explained the text... and then new grammar... all in Chinese, of course. After that, we read aloud the text, the whole class together. For homework... we often do [did] multiple-choice exercises, normally 100 MC questions... every day. On the second morning, the teacher checked the answers to the MC questions... and then gave us dictations to see if we have [had] remembered the new words in the lesson....

Writing? No, we had no writing exercises... until one month before our college entrance examination.

Vocabulary Learning

The evidence of intentional learning was most manifest in the respondents' lexical learning experiences. Most participants in the interviews reported that the ability to recognize the form of a word and retrieve its meaning in Chinese was the most important aspect of their lexical learning. They mainly learned words from word lists. These lists could be from their own textbooks and teachers' notes, as well as from NMET and TOEFL (Test of English as a Foreign Language) preparation books. A key factor leading to the popularity of these lists was that they contain complex lexis that are regarded as "high-level" words by many students. Learning complex words, according to many participants, was crucial for achieving a high NMET score in the writing section. The reason was explained by a participant:

An important criteria [criterion] to judge our writing... in the college entrance examination is... you should have some... so-called... high-level words in your writing. I mean, if you can use "significance," you should not use "importance." You should use words... words that look longer and more complicated and... something like that. It works!... It give [gives] more marks to your writing.

With such beliefs on vocabulary learning, some respondents memorized as many as 100 new words per day, which was achieved mainly via intentional learning. Their teachers helped them by delivering frequent dictations in class. These findings added support to the observations of Zhao (2016).

Training in Writing

Given that their English learning in secondary school was dominated by intentional learning, which focused largely on vocabulary and grammar training, most participants reported that they had very limited experience in writing. The most common experience was that their training in writing did not start until the last few months of their secondary school education, to be exact, one to three months before the NMET. The so-called training was essentially test-taking strategy training. Due to the task requirements, task prompts, and adopted assessment criteria of the writing task on the NMET, the students at the interviews reported that testwiseness strategies (Cohen, 2006) were mainly adopted to train them to manage the writing task. These strategies enabled them to take advantage of “the characteristics and formats of the test” to attain a high grade (Millman et al., 1965, p. 707). To achieve this purpose, teachers asked their students to attempt writing tasks from past NMET papers and mock-test papers and then trained them in the skills necessary to manage similar tasks. This finding accords with the observations of Kennedy and Lui (2013).

According to the participants, an overarching area accentuated by most teachers during the training was the discrete-point knowledge of English grammar and vocabulary. Most teachers believed these two facets of knowledge were the cornerstones of effective writing. Another area that was centered on was students’ ability to use complex sentences and “beautiful structures” (You, 2010), also known as “shining phrases” by the participants. Examples of these phrases and complex sentences that the teachers expected students to use in examinations can be seen in a predetermined composition template provided by a respondent at the interviews. Memorizing composition templates was one of the key strategies used by these learners to obtain a high score on the NMET.

It is well known to us that the proverb “_____ (state the proverb)” has **profound** significance and value not only in our job but also in our study. It means _____ (explain the proverb). The saying can be **illustrated** through a series of examples as follows. A case in point is _____ (give your example). Another case is _____ (give your second example). Therefore, it goes without saying that it is of great importance to practice proverb _____ (restate the proverb). With this rapid **development** of _____ (state the essay topic), an increasing number of people come to realize that it is _____ (link

the proverb to the essay topic). The more we are aware of the significance of this famous saying, the more benefits we will get in our daily job and life.

This template is for writing expository essays. The italicized parts (in parentheses) were in Chinese in the original template and were translated into English for the purpose of the current study. In this template, there are altogether eight sentences, of which seven are complex sentences (only the third one is a simple sentence). These eight sentences contain six complex words (in bold) and seven “beautiful structures” / “shining phrases” (underlined).

Since complex words/sentences and “beautiful structures” play such an important role in helping students achieve high grades on the NMET, most students worked diligently on these items. A corollary of this practice was that many students’ writing had become cumbersome and labored, which could cause difficulties in readers’ comprehension of the written texts. This problem can be seen more clearly in the two case studies. The two cases are Jacky, from a small city, and Mary, from a large city. The two cases represent learners in cities at different stages of development in China. Due to the researcher’s close interaction with Jacky and Mary during the process, the personal pronoun “I” was used in the presentation of the two cases.

Case Studies

Case I—Jacky

Jacky was a student from a small city in China. He was in one of my seminar groups for an EAP course. He came to my attention due to his unusual passivity in class. He seldom spoke up in class. Whenever asked to answer a question, he could hardly utter a complete sentence for his answer. He received a D for his first in-class written assignment, a grade markedly low in his group. His situation captured my attention from the very beginning of the semester. I, therefore, observed him carefully and communicated with him on a number of occasions with the intention of helping him perform better on the subject.

From our conversations, I learned that his passivity in class was caused by his low confidence in his productive language skills. He paid little attention to his speaking skills since speaking was not tested on the NMET. His training in writing did not formally start until a few months before he took the NMET. Nevertheless, with diligence and perseverance,

Jacky attained a very high score on the NMET. His overall score was 136 (out of 150), and his score on the writing task was 26 (out of 30). He attributed his success to his large vocabulary, his high command of English grammar, and his teacher's strategic training to help him manage the writing task on the NMET.

Jacky's teacher demanded students' particular attention to two factors in writing: tidy handwriting and "high-level" words. According to the teacher, inasmuch as a composition satisfied these two criteria, even if sentences in the composition made little sense of the given composition topic, a mark of at least 21 (out of 30) could be given. His teacher also trained him to memorize exemplary essays for the NMET and, in particular, "high-level" words and formulaic expressions in these essays. By the time Jacky took the NMET, he had memorized at least 500 such sentences. These sentences were for different parts of an essay; that is, some were specifically for writing introductions, some for conclusions, and others for body paragraphs. Below are examples of sentences that Jacky memorized for an argumentative essay:

Sentences for an introduction:

1. Recently, the problem of... has aroused people's concern.
2. There are different opinions among people as to.... Some people suggest that....
3. Everything has two sides and... is not an exception, it has both advantages and disadvantages.

Sentences for a conclusion:

1. There is no doubt that... has its drawbacks as well as merits.
2. Taking into account all these factors, we may reasonably come to the conclusion that....
3. Hence/Therefore, we'd better come to the conclusion that....

Equipped with these sentences, Jacky needed to activate only some of them, changing words here and there based on the composition topic, and then added a few other sentences to link the memorized sentences together to achieve a high grade on the writing task of the NMET. However, his lack of incidental learning backfired when he came to the university, where the medium of instruction is English. With my assistance and his assiduous effort, Jacky's speaking and writing skills showed some improvement at the end of the semester. Nevertheless, his journey to becoming a successful language learner would still be a long and onerous one. This is evidenced by a short paragraph taken from the

out-of-class essay he wrote for the last assessment of the semester. This paragraph showed lexical errors of various kinds in their language production, particularly in word choices, word collocation, and lexicogrammar:

As well as the unfair in education, the wealth gap may also lead to social instability. The poor people would resentment the wealthy people and may even abuse violent to protest against the social unfair. According to a research, 96% publics hate the wealthy and myriad feel discontented toward the government. In 2011, a plethora of people protest the low income and high cost violently.

This essay was about the impact of an increasing wealth gap in Hong Kong. What he intended to argue was that the growing wealth gap causes much resentment among the poor toward the rich. This resentment may also extend to the government and all of society. To express discontent with their worsening situation, many poor people may stage mass protests, sometimes leading to violence. However, this intended argument was not very clearly presented, largely due to the lexical errors in the paragraph (underlined). These errors are mainly in the following five categories (intended words are italicized and placed in parentheses).

- a. word choices: would (*could*), abuse (*use*), and research (*study*)
- b. morphological forms: unfair (*unfairness*), resentment (*resent*), and violent (*violence*), publics (*public*)
- c. word collocation: abuse violent (*use violence*), discontented toward (*with*), and protest... (*protested against...*)
- d. Lexicogrammar: would (*may*), hate (*hated*), and myriad (*the myriad*)
- e. unnecessary use of complex (low frequency) words: myriad (*a countless or extremely great number of*) and plethora (*a large or excessive amount of*)

The above lexical errors suggest that Jacky may have a large vocabulary that he gained through intentional learning, but his lexical competence, the ability to use lexis appropriately in communication, was still rather limited. The analysis also indicates the centrality of interdependence between different parts of language skills (Ellis, 2014). In other words, dependence on intentional lexical learning can affect learners' lexicogrammar and other writing skills.

Case II—Mary

Mary was from Beijing, one of the most developed cities in China. Similar to Jacky, Mary attained a very high grade on the NMET (142 out of 150), particularly in writing (27 out of 30). Despite this high writing score, Mary's teacher, who was a native speaker of English, found it difficult to understand her writing and asked me to follow up with the case due to my role as the subject leader. During my several meetings with her, I found that Mary mainly attended to two items that she believed were essential in English learning: memorizing as many low-frequency words as possible and composing long, complex sentences in the EW portion. To achieve the vocabulary goal, she memorized words and phrases in her textbooks as well as in TOEFL and even GRE (Graduate Record Examinations) preparation books. To accomplish her mission of using long and complex sentences in writing, she diligently worked on grammar and used many sentence connectors (mainly conjunctions) in her essays.

Mary's strategies seemed to have worked well, considering her outstanding performance on the NMET. Nevertheless, she was still very concerned about her examination skills and asked me many questions in this regard during the meetings. One of the consequences of her focus on test-strategy training was her reluctance to spend time on any incidental learning. To her, activities such as reading and listening "take too much time" and thus were not worth trying. Whenever she engaged herself in independent language learning, she memorized complex words in her word lists or read an examination-preparation book. The consequence of such training was her tense and ponderous style of writing, which was full of low-frequency words and overly complex sentences but often made little sense to readers.

With my assistance and possibly also other teachers in the university, her writing showed some improvement by the end of the semester, but the progress was "very limited," according to the instructor of her class. Mary claimed this was because she did not even know "how to write simple sentences anymore." Below are two unnecessarily long and overly complex sentences extracted from one of Mary's essays. The essay is about the balance between press freedom and the protection of privacy.

1. Society should attach more importance to the freedom of expressing opinions and comments reasonably by the media rather than the

freedom of unveiling the private lives of citizens, which is regarded as offensive and impolite.

2. Furthermore, exposing the private lives of celebrities to the public for entertainment purposes is actually barely constructive but only damages the reputation of the victims and raises concerns about privacy, which are significantly threatened.

There is no obvious sentence structure error in these two sentences. However, her complex sentence structures and inaccurate word use make it difficult for readers to fully capture the meaning of the sentences. These sentences provide evidence for the long-term negative impact of intentional learning on Mary's writing skills. Yes, she may have gained a large vocabulary. With extensively trained test-taking skills, she may also have received a high score on DVK. However, her lexical skills, in particular the facets of automaticity and appropriateness of word knowledge, were still underdeveloped.

The above results from the three focus group interviews and two case studies have provided the answer to RQ3, that is, a singular lack of incidental vocabulary learning has severely impacted learners' writing experience and proficiency. These results shed light on the weak association between learners' performance on DVK and EW and the insignificant 1% change that DVK added to the variance already explained by VS. With test-taking skills, these learners were able to achieve reasonable scores on DVK, but the interconnectedness between words in their mental lexicon, which can be achieved mainly via incidental learning, was still rather underdeveloped. The results showed that dependence on intentional learning limited the learners' opportunities to write in English, which in turn affected their development of effective writing skills. The cases of Jacky and Mary also demonstrated that lacking incidental learning undermined the learners' confidence in writing and caused their misconceptions of what counts as good writing. Their low confidence and misconception further hindered the development of the learners' writing skills.

Conclusions and Implications

This paper examines the role of vocabulary knowledge in developing EFL learners' writing skills. The empirical evidence from this study has further underscored the centrality of productive vocabulary knowledge in enhancing learners' writing experience and developing their writing

proficiency. The study, however, is not without limitations. The major limitation lies in the fact that almost all the participants were graduates of elite secondary school in Mainland China. If other demographics of learners, particularly those whose NMET scores do not satisfy the English proficiency requirements of Hong Kong universities, are investigated, the impact of the lack of productive lexical knowledge on writing skills should be even more evident. Notwithstanding the limitation, this multifaceted investigation, via (inter alia) analyses of the association between VS, DVK, and EW and the three focus group interviews, has drawn a broad-brush landscape of learners whose writing skills were negatively impacted by a paucity of productive vocabulary knowledge. The two case studies have also painted portraits of individuals who suffered from the severe impact.

The study has two significant pedagogical implications. The first is the centrality of conducive contextual language environments for learners, that is, offering learners sufficient and high contextual exposure to the target language. Achieving this goal requires the concerted efforts of different stakeholders in the education system. Efforts in this regard could perhaps start by reforming the relevant examination systems. Test designers for the MNET, for example, could consider placing the test focus more on learners' communication skills rather than discrete-point lexical and grammar items. Such changes may facilitate positive backwash (Atmojo, 2021) so that learners will pay more attention to language production skills in their learning. This shift in attention may increase the need for authentic reading and listening materials, which can then help improve the contextual factors in China.

The second implication is a balanced provision for intentional and incidental learning in learners' lexical development. The most effective approach to achieving this goal is perhaps to ensure the adoption of Nation's (2007) four strands framework in schools. This milestone framework allows a balance between *language-focused learning* and *meaning-focused learning*. Language-focused activities via intentional learning develop the foundation and some specific aspects of lexical knowledge for learners. Meaning-focused activities through incidental learning complement the former by offering repeated encounters with input over time. The combination of these two lexical learning strategies can effectively develop the interrelated connections between different dimensions of word knowledge in learners' lexicons to facilitate their authentic communication skills.

References

- Atmojo, A. E. P. (2021). Facilitating positive washback through authentic assessment in EFL assessment. *Journal of Teaching and Education*, 3(2), 226–233.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Butler, Y. G. (2014). Parental factors and early English education as a foreign language: A case study in Mainland China. *Research Papers in Education*, 29(4), 410–437.
- Cohen, A. D. (2006). The coming of age of research on test-taking strategies. *Language Assessment Quarterly*, 3(4), 307–331.
- Crossley, S. A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3), 415–443.
- Ellis, N. C. (2014). Cognitive and social language usage. *Studies in Second Language Acquisition*, 36(3), 397–402.
- Færch, C., Haastруп, K., & Phillipson, R. (1984). *Learner language and language learning*. Multilingual Matters.
- Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. University of Michigan Press.
- Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Taylor & Francis.
- Gu, P. Y. (2020). Strategies for learning vocabulary. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 271–287). Routledge.
- Halliday, M. A. K., & Matthiessen, C. M. (2013). *Halliday's introduction to functional grammar*. Routledge.
- Hamp-Lyons, L., & Kroll, B. (1996). Issues in ESL writing assessment: An overview. *College ESL*, 6(1), 52–72.
- Hu, G. W. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development*, 24(4), 290–318.
- Jin, L., & Cortazzi, M. (Eds.). (2011). *Researching Chinese learners: Skills, perceptions and intercultural adaptations*. Springer.
- Kennedy, S., & Lui, R. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concordia Working Papers in Applied Linguistics*, 4, 22–29.
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), 365–391.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.

- Lin, L. H. F. (2019). The NMET impact on the English writing of Mainland Chinese students. *ERL Journal*, 1(1), 122–131.
- Lin, L. H. F., & Morrison, B. (2021). Challenges in academic writing: Perspectives of engineering faculty and L2 postgraduate research students. *English for Specific Purposes*, 63, 59–70.
- Liu, M., & Jackson, J. (2011). Reticence and anxiety in oral English lessons: A case study in China. In L. Jin & M. Cortazzi (Eds.), *Researching Chinese learners: Skills, perceptions, and intercultural adaptation* (pp. 119–137). Palgrave Macmillan
- Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33–42.
- Llach, M. P. (2011). *Lexical errors and accuracy in foreign language writing*. Bristol: Multilingual Matters.
- McCarthy, M. J. (1990). *Vocabulary*. Oxford University Press.
- Meara, P. M., & Alcoy, J. C. O. (2010). Words as species: An alternative approach to estimating productive vocabulary size. *Reading in a Foreign Language*, 22(1), 222–236.
- Millman, J., Bishop, C. H., & Ebel, R. (1965). An analysis of test-wiseness. *Educational and Psychological Measurement*, 25(3), 707–726.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 1–12.
- Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 125–160). Multilingual Matters.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge. *Canadian Modern Language Review*, 56(2), 282–308.
- 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. (1998). Validating a test to measure depth of vocabulary knowledge. In A. J. Kunnan (Ed.), *Validation in language assessment: Selected papers from 17th Language Testing Research Colloquium, Long Beach* (pp. 41–60). Lawrence Erlbaum Associates Publishers.
- Reid, J. (1993). *Teaching ESL writing*. New York: Prentice Hall.
- Saldaña, J. (2021). *The coding manual for qualitative researchers*. SAGE Publications Ltd.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951.

- Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2017). How much vocabulary is needed to use English? Replication of Van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, 50(2), 212–226.
- Schmitt, N., Dornyei, Z., Adolphs, S., & Durrow, V. (2004). Knowledge and acquisition of formulaic sequences: A longitudinal study. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 55–86). John Benjamins.
- Shaw, S., & Falvey, P. (2008). *The IELTS Writing Assessment Revision Project: Towards a revised rating scale*. University of Cambridge ESOL Examinations.
- Singleton, D. (2007). How integrated is the integrated mental lexicon? In Z. Lengyel & J. Navracscics (Eds.), *Second language lexical processes: Applied linguistic and psycholinguistic perspectives* (pp. 3–16). Multilingual Matters.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139–152.
- 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.
- Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). Routledge.
- Yanagisawa, A., & Webb, S. (2020). Measuring depth of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 371–386). Routledge.
- You, X. (2010). *Writing in the devil's tongue: A history of English composition in China*. Carbondale: Southern Illinois University Press.
- Yu, J., Brown, G. T., & Stephens, J. M. (2018). Retrospective case studies of successful Chinese learners of English: Continuity and change in self-identities over time and across contexts. *System*, 72, 124–138.
- Zhao, J. (2016). The reform of the National Matriculation English Test and its impact on the future of English in China: Will English lose its predominance in the Chinese foreign language landscape? *English Today*, 32(2), 38–44.

Linda H. F. LIN holds a Ph.D. in applied linguistics and works at the Hong Kong Polytechnic University. She has extensive experience in L2 language learning and teaching, L2 language writing analysis, and pedagogic development. Her research interests include academic writing, vocabulary learning, and applications of concordancing in teaching and learning. She has published articles in journals such as English for Special Purposes, English for Academic Purposes, and CALICO Journal, as well as in books such as The Routledge Handbook of Vocabulary Studies. She is also an editor of the ELR Journal.