課程與教學季刊　2014, 17(1),頁 95～124

The Relationship of EFL Learners' Vocabulary Knowledge to Their Speaking Performance: Productive Vocabulary Size, Collocational Knowledge and Idiomatic Knowledge

Szu-Ling Tseng *　Wen-Ying Lin **

This purpose of the study was to investigate the relationship of Taiwanese learners' English speaking performance (ESP) to their performance on each of three vocabulary knowledge subcomponents. Specifically, the study aimed at answering the following three research questions: (1) What is the relationship of ESP to productive vocabulary size (PVS)? (2) What is the relationship of ESP to lexical collocational knowledge (LCK) and to actual use of lexical collocational knowledge (AULCK)? (3) What is the relationship of ESP to idiomatic knowledge (IK) and to actual use of idiomatic knowledge (AUIK)? A total of 32 graduate students from two universities in northern Taiwan participated in the current study. Five instruments were employed, including (1) one productive vocabulary levels test, (2) one lexical collocation test, (3) one idiom test, and (4) two speaking tests. The collected data were analyzed through Pearson correlation procedures. The results of the present study showed that ESP was significantly and moderately related to PVS and to IK. However, no significant correlation was found between ESP and LCK. Neither was there any significant relationship found between ESP and AULCK, and between ESP and AUIK. Based on these findings, some implications and suggestions for future research were provided.

Keywords: English speaking performance, productive vocabulary size, lexical collocational knowledge, idiomatic knowledge

*　Szu-Ling Tseng, English Teacher, Touhu Elementary School, New Taipei City

**　Wen-Ying Lin, Associate Professor, Department of English Instruction, University of Taipei

Corresponding Author:Wen-Ying Lin，e-mail: wylin@utaipei.edu.tw

Motivation

In recent years, English speaking skills have been considered more and more important in Taiwan, where English is taught and learned as a foreign language (EFL). One of the reasons for receiving increasing attention may come from a requirement recently set by most universities. That is, many universities mandate that prior to graduation, students should pass at least one well-established standardized English proficiency tests, such as TOEIC (Test of English for International Communication), TOEFL (Test of English as a Foreign Language), IELTS (International English Language Testing System), and GEPT (General English Proficiency Test). As most of these tests include a speaking subtest, making it seem all the more urgent for English teachers at the tertiary level to place strong emphasis on English speaking classroom instructions. For example, in recent years proficiency in English speaking has become one of important objectives for university English curriculum in Taiwan.

Important as it is, English speaking proficiency is not easy to achieve and often poses a great challenge for most students in Taiwan. Just as numerous researchers (i.e., Berman & Cheng, 2000; Iwashita, Brown, McNamara, & O'Hagan, 2008) stated, EFL speaking proficiency is probably the most difficult target to achieve, for it is perceived as a productive skill involving a combination of various knowledge or competences, such as accurate pronunciation, grammatical competence, vocabulary knowledge, rhetorical organization, and so on. Therefore, despite the fact that most Taiwanese college students have started to learn English since they were studying in elementary schools or even kindergartens, most of them are still unable to fully and fluently express their thoughts. What's worse, some of them can hardly produce a complete and correct sentence in a natural English conversation (Chuang, 2010). In fact, according to the record released by the Language Training and Testing Center in Taiwan ("Score Data Summary for 2012 GEPT," 2012), the 2012 passing rate of High-Intermediate GEPT subtests on productive English skills (i.e., speaking and writing skills) was only 25%, which was much lower than 44%, the 2012 passing rate of its subsets on receptive English skills (i.e., listening and reading). Hence, in spite of their early exposure to English learning, speaking proficient English seems to be a mission impossible for many learners in Taiwan.

Among the various knowledge or competences that are perceived to affect

English speaking performance (hereafter ESP), English vocabulary knowledge, which has long been perceived to be a multi-dimensional construct, has been considered by many researchers (e.g., Read, 2000) as playing a significant role in ESP. Just as Wilkins (1972, p.111) stated, "Without grammar very little can be conveyed, without vocabulary nothing can be conveyed." From this statement, it is not difficult to see that English vocabulary knowledge is definitely an important factor contributing to effective ESP. Such a claim is also supported by many empirical studies (e.g., Iwashita et al., 2008). For example, investigating the distinguishing features of TOEFL iBT (Internet-based test) speaking performance, Iwashita and her colleagues (2008) found that of all the features, vocabulary use was one of the main factors affecting 200 students' overall speaking scores across five levels of general English proficiency.

Given the empirically evidenced importance of English vocabulary knowledge to ESP, the present study was called for in an attempt to take a close look at the relationship between the various dimensions of English vocabulary knowledge and ESP. It was hoped that the results of the present study could guide English language instructors and teaching material designers toward pedagogically sound practices with respect to the learning of both English vocabulary and English speaking.

Literature Review

In light of the fact that English vocabulary knowledge is essential for effective ESP, it is worth knowing what components constitute speakers' knowledge of words. According to Anderson and Freebody (1981), vocabulary knowledge can be classified into two dimensions: breadth and depth of vocabulary knowledge. Breadth of vocabulary knowledge or vocabulary size refers to how many words a learner knows, while depth of vocabulary knowledge or quality of vocabulary knowledge refers to how well s/he knows a word. To put it more comprehensively, knowing a word includes the knowledge of "not only its semantic features but also its orthographic, phonological, morphological, syntactic, collocational and pragmatic characteristics" (Read, 2004, p. 211). Both breadth and depth of vocabulary knowledge can be further classified into receptive and productive vocabulary knowledge (Nation, 2001). The former pertains to the ability to recognize the meaning of a word in listening or reading, whereas the latter refers to the ability to produce and use a word in speaking or writing.

To date, abundant empirical evidence from western studies (e.g., Koizumi, 2005; Zimmerman, 2004) has been obtained to support the claim that vocabulary size plays a crucial role in learners' ESP. For example, the relationship between productive vocabulary size (hereafter PVS) and ESP has been examined by Zimmerman (2004) in USA and by Koizumi (2005) in Japan. In Zimmerman's study, a moderate correlation coefficient (r = .66, p < .0001) was found between 173 college students' ESP and their PVS. Likewise, Koizumi (2005) reported a positive and moderate to high correlation coefficient (r = .77, p < .01) between 225 high school students' ESP and PVS. These findings suggest that learners with a big PVS tend to speak English more proficiency than learners with a small PVS.

Although empirical support was obtained from numerous overseas studies for the significant relationship of ESP to PVS, very few studies can be found in Taiwan to examine these two variables together. Most of Taiwanese empirical studies on vocabulary size aimed to examine its relation to L2 general language proficiency (e.g., Liu, 2002), reading (e.g., Hsu, 2009), and listening (e.g., Tsai, 2005). Hence, a need is warranted to conduct a study to investigate the relationship of Taiwanese learners' ESP to PVS.

In addition to PVS (i.e., breadth of vocabulary knowledge), depth of vocabulary knowledge, especially collocational knowledge, has also been claimed to be related to EFL learners' ESP (Shin, 2007). Collocational knowledge, according to Ellis (2001), is learners' knowledge about how likely words occur with others and their ability to store chunks of language in long-term memory. Numerous researchers (e.g., Granger, 1998) have posited that collocational knowledge enables learners to reduce cognitive effort, save processing time, have ready-made chunks available, and make more native-like speaking performance. If EFL learners use many appropriate collocations while taking a speaking test, raters may consider them as fluent speakers. On the other hand, the raters' perception of their performance may become negative when they use few collocations while speaking. The relationship between collocational knowledge and ESP is claimed to be strong.

Unfortunately, this claim has been empirically supported only by few studies in western countries. The first study was done by Sung (2003) to examine lexical collocational knowledge (hereafter LCK), as measured by Test of Lexical Collocations. The study also tried to assess actual use of lexical collocational knowledge (hereafter AULCK) by counting the number of lexical collocations actually used by a total of 72 undergraduate and graduate students in their oral

descriptions and feedback about a film watched. A significant relationship was found both between their scores of LCK and their ESP (r = .78, p < .01), and between their' AULCK and their ESP (r = .55, p < .01). Another similar study, conducted by Ushigusa in 2008 only on AULCK of 38 graduate students who were non-English majors, also reported a significant but smaller relationship (r = .46, p < .05) to ESP.

Likewise in Taiwan, not too many studies have been conducted along this line. The first attempt to explore this issue in Taiwan was made by Hsu (2007). Eleven college students' ESP in an impromptu speech contest was recorded, rated, and analyzed. His analysis focused on the relationship between the contestants' ESP and AULCK, and a significantly moderate relationship (r = .59, p < .05) was obtained. However, his results may have to be interpreted with some caution since his contestants' ESP in an impromptu speech contest may not truly reflect their English speaking behavior in their daily lives, which was expected to be more natural, authentic and less nervous. Furthermore, unlike Sung's study, LCK was not assessed with a pencil-and-paper test. Instead, only the contestants' AULCK was examined in Hsu's study. In response to the two limitations, Hsu further explored this topic with his graduate student one year later. This study (Chiu & Hsu, 2008), following Sung's study (2003), adopted two speaking tests that were more authentic in nature and assessed a wider range of ESP than the one used in his 2007 study. Furthermore, in addition to counting the number of lexical collocations actually used by 56 college students in the Pear Film speaking test (one of his two speaking tests), this study also included a pencil-and-paper measure of LCK. A significant and moderate correlation coefficient (r = .56, p <.01) was found between LCK and overall ESP. However, contradictory to the results of Sung's (2003), no significant correlation (r = .25, p > .05) was found between the learners' AULCK and their ESP. Therefore, it still remains inconclusive about the relationship between AULCK and ESP. In addition, the participants of the above studies were non-English majors. Therefore, it still remains unknown about whether their results can be generalized to Taiwanese English-major graduate students, who may tend to be more proficient in English than those non-English majors. That is, the relationship between LCK and ESP performance awaits further investigation with learners of relatively high English proficiency levels.

In addition to LCK, idiomatic knowledge (hereafter IK) is also an important factor that may affect ESP. It is referred to as the understanding of "relatively invariable expressions with meanings that cannot be predicted from the meanings of

the parts" (Biber, Johansson, Leech, Conrad, & Finnegan, 1999, p. 988). According to John-Laird (1993), it is quite hard to "speak spontaneously without lapsing into idiomatic usage" (p.3). McDevitt (1993) also stated that "idioms are an important part of any language and may be said to be an indicator of one's fluency in that language" (p.4). Similarly, Ellis (1997) also argued that fluent use of IK is "an important index of native-like competence" (p. 130). Using appropriate IK allows ESL/EFL learners to speak like English native speakers who store idioms as figurative meanings in their memory. In other words, without IK, ESL/EFL learners' ESP tends to be judged as unnatural, odd, or foreign although their grammar is correct. However, such a claim was supported by only one empirical western study (Ushigusa, 2008), where a significant and moderate correlation coefficient (r = .53, p < .05) was obtained between actual use of IK (hereafter AUIK) and ESP among a total of 38 Chinese speaking learners. Likewise in Taiwan context, only one attempt (Hsu, 2007) has been made to investigate the relationship between the AUIK and ESP. Specifically, besides LCK, Hsu (2007) also intended to find out the relationship of ESP to AUIK. However, it turned out that no correlation coefficient was reported because none of the 11 contestants actually used any idioms in the speech contest. Hence, more studies are definitely needed to verify whether AUIK or IK is related to ESP.

Another point worth mentioning is that most studies assessing ESP tended to have some limitations with respect to the validity or appropriateness of their measures. Take Chiu and Hsu's (2008) study for example. In addition to the Pear Film speaking test, PhonePass spoken test was also used in the study to measure their test takers' ESP over the telephone by combining computerized design, telephone, and the Internet. It included four parts: reading, repeating, answering short questions, and building sentences. In the reading section, the students had to read sentences shown in test sheets, and in the repeating part, they were asked to repeat what they just heard on the telephone. The two subtasks aimed to measure EFL learners' pronunciation, reading and listening skills, not their self-utterance of spoken production. Furthermore, the test was conducted individually at home, and thus there was no guarantee that the test was really taken by the participants. Hence, the test might not be quite valid for assessing ESP. In addition, since the test takers were required to talk over the phone and recording of their telephone speaking performance was not available, the test takers' AULCK could not be counted. Thus, as recommended by Chiu and Hsu (2008) in their conclusions, future studies should include various speaking test tasks, such as face-to-face conversations, interviews,

or opinion/experience sharing to elicit test takers' ESP.

Taken as a whole, a large body of research has been conducted on the relationship of EFL learners' ESP to PVS (e.g., Koizumi, 2005; Zimmerman, 2004), LCK (e.g., Chiu & Hsu, 2008; Hsu, 2007; Sung, 2003), and IK (e.g., Hsu, 2007; Ushigusa, 2008). However, no research has investigated the three components all together in a single study. Therefore, it seems that there is a need to conduct a study along this line. Additionally, as commented by Chiu and Hsu (2008), each testing technique employed to assess ESP is bound to have its limitations in terms of validity or appropriateness. Hence, future studies that involve the assessment of ESP should try to incorporate a variety of testing tasks, so that a wide spectrum of the construct can be elicited and evaluated.

Research Questions

The purpose of the present study is to investigate the relationship of ESP to PVS, LCK, and IK. Specifically, the research questions of this study are as follows:

1. What is the relationship between PVS and ESP?

2. What is the relationship between LCK and ESP?

3. What is the relationship between IK and ESP?

Method

Participants

The participants of the study were a total of 32 graduate students (6 male and 26 female) majoring in English at two universities in northern Taiwan. Most of them were in their first, second or third year of graduate study with their ages ranging from 23 to 34 years old. All of them had received formal EFL instruction for more than 10 years, and had either passed High-Intermediate GEPT or obtained a score above 750 for TOEIC.

Instruments

Five instruments were used in the study, including a productive vocabulary levels test, a lexical collocation test, an idiom test, and two speaking tests. Each of them was described in detail in the following paragraphs.

The Productive Vocabulary Levels Test (PVLT). Due to its ease of administration and scoring, the Productive Vocabulary Levels Test (hereafter, PVLT) was chosen and used in the present study as a measure of the participants' PVS. Developed by Laufer and Nation (1999), it was a paper-and-pencil test containing five levels of word frequency: the 2,000 word level, the 3,000 word level, the 5,000 word level, the university word list level, and the 10,000 word level. Due to the time constraint, the present study tended not to test all the vocabulary levels. Instead, it included only 2,000, and 3,000 word levels, since most Taiwanese college students' average PVS has been reported as near or less than 2,000-3,000 words (Cheng, 2007).

For both word levels of the test, four versions (A through D) were available. As Version C was readily accessible from an appendix of Laufer and Nation's (1999) paper, it was adopted in the present study. A high reliability estimate of about .91 on KR21 was reported. The items were in the format of completion, requiring the test takers to complete the underlined words. For each item, a sentence was given to provide the context and the first few letters were given to cue the tested word. By doing so, test takers could produce the words rather than simply recognize them. There were 18 questions for each of the two word levels, and thus the whole test had 36 items, each of which was worth one point. As such, the maximum possible score was 36 points. An example of the items is: I'm glad we had this opp_____ to talk.

Lexical Collocation Test (LCT). The fill-in-the-blank Lexical Collocation Test (hereafter LCT), mainly adapted from the one developed by Chiu and Hsu (2008), was used in the present study to measure the participants' LCK. Instead of including all seven types of lexical collocation categorized by Benson, Benson and Ilson (1997), Chiu and Hsu incorporated only five categories of lexical collocations in their test: Verb + Noun, Adjective + Noun, Noun + Verb, Adverb + Adjective, and Verb + Adverb (target words are underlined), with each category containing 10 test items. Specifically, Benson et al. (1997) first two types were combined into one category in Chiu and Hsu's test because they shared the same pattern: verb + noun. Moreover, Benson et al. (1997) fifth type (noun1 of noun2) was disregarded, as it was somehow controversial and thus was not considered a lexical collocation by some studies (e.g., Hsu, 2006). Based on the list of 1,000 Basic English Words for

Junior High School and Elementary School Students, published by the Department of Ministry of Education (MOE) in 2003, a total of 50 words were selected by Chiu and Hsu for eliciting test takers' target collocations. Each of the 50 selected words occurred in the context with at least two to three sentences, so that there was enough information for the test takers to retrieve the correct collocation. However, unlike the test developed by Chiu and Hsu (2008), the test used in the present study did not provide the first letter or phoneme of the answer to each item, for the purpose of avoiding giving obvious clues. The following is a sample item:

It is polite to _____ hands at the end of a business meeting when the host and guest need to leave. This implies friendliness, trust and the possible lead to future negotiations. However, this is not appropriate in all cultures. Investigate local customs if you will be visiting a foreign country. Answer: *shake

For ease of comparison, the scoring method used by Chiu and Hsu (2008) was also adopted in the present study. Specifically, a three-point partial-credit scoring method suggested by Aghbar and Tang (1991) were used to rate the responses on the LCT. Every item in each of the five parts was assigned zero to two points. Thus the maximum possible total score was 100 points for the 50-item test. The accuracy of the answers to the questions was evaluated, based on the following references: *BBI dictionary of English word combinations* (Benson, et al., 1997) and *Oxford Collocations Dictionary for Learners of English* (Lea, 2002). Additionally, two free online corpora, Corpus of Contemporary American English (http://sara.natcorp.ox.ac.uk/lookup.htm) and VLC Web Concordancer (http://corpus.byu.edu/coca/), were also consulted because numerous studies showed that a corpus could comprehensively illustrate how words collocate (e.g., McCarthy, 2004). However, following Chiu and Hsu (2008), the present study likewise awarded just one point to a collocate response that could only be found in either *BBI dictionary of English word combination* or *Oxford Collocations Dictionary for Learners of English*, as online corpora may include many possible collocations which are used loosely in different registers and genres. Moreover, grammatical errors and misspelling were neglected since the purpose of this test was to assess the participants' lexical collocational knowledge rather than their grammar or spelling. The reliability estimates for scores on the LCT was .72.

Test of Idioms (TOI). Adapted from Liu's (2006) idiom test, Test of Idioms (hereafter TOI) used in the current study aimed at assessing the test takers' IK. Like

Liu's test, TOI was also made up of 25 items, each containing one idiom. However, only 13 out of the 25 idioms in Liu's test were adopted in the current study, as the remaining 12 idioms of his test were deemed to be way too difficult by two EFL scholars. As such, 12 more idioms of the TOI were selected from various commonly-used senior high school English textbooks published by Far East, Sanmin, Nani, and Lungten publishers. Two criteria for idiom selection were set: each of the idioms selected from the textbooks could also be found in *Oxford Learner's Dictionary of English Idioms* (McCatig, 1994), and words that formed the idioms could all be found in the MOE 1,000 Basic English Wordlist. In order to use the idioms in a productive way, the participants were required to construct and write a meaningful sentence or sentences for each idiom. One example of the items is presented in the following:

Direction: Please make a meaningful sentence/sentences for each idiom.

1. around the corner

Following Liu's (2006) three-point rating scheme (see Table 1), each answer was assigned a score ranging from zero to two points, with a possible maximum score of 50 points. Two native speakers were invited as raters in the current study. Both of the native speakers had been teaching English in colleges or universities in Taiwan for more than 10 years in Taiwan. The first native speaker had a M. A. degree in TESOL in San Francisco State University. The other rater held a M. A. degree in English Literature in University of Notre Dame. Prior to the formal administration of the test, a training session was held to ensure inter-rater reliability. At the beginning of the training session, the purpose of the study and the scoring methods were introduced first, so that the two raters could fully understand the rating criteria. Trained with sound judgment in scoring, the two raters were then asked to rate two test sheets selected at random for practice. After all the scoring was done separately and independently, the two sets of scores were compared and discussed for the purpose of making sure that both raters reach an agreement on the scoring. Discrepancies in the scoring between the raters were settled through discussion. The inter-rater reliability estimate for scores on the TOI was .94.

Table 1 *The Three-Point Criteria Used for the TOI*

Level	Scoring Criteria	Points
Level 3	Grammatically, semantically, and contextually correct. Ex. He has a sweet tooth because he can't stop eating dessert.	2 points
Level 2	Grammatically, semantically, and/or contextually ambiguous. Ex. I have a sweet tooth.	1 point
Level 1	Grammatically, semantically, and/or contextually incorrect. Ex. He has a sweet tooth because he likes to say something sweet.	0 point

The Speaking Tests. In the present study, the participants' ESP was assessed by the following three indicators: the Pear Film speaking test, a published sample speaking test taken from IELTS 7 (2009), and the average of scores of the two speaking tests. The two speaking tests were chosen because of the following reasons. First of all, Pear Film speaking test, which was also used in Chiu and Hsu's (2008) and Sung's (2003), was employed in the present study to facilitate comparison of results among studies. Second, unlike the PhonePass speaking test (used in Chiu and Hsu's study in 2008), where recording of speaking behavior was not feasible, the participants' oral responses in the both tests of the present study could be recorded, coded, and analyzed. Finally, the participants' speaking behavior could be sampled and evaluated through two different testing techniques---question and answer from the Pear Film speaking test, and interview from the IELTS speaking sample test; hence, the construct validity of ESP and the generalizability of the results could be enhanced. Each of the two tests is described in the following paragraphs.

Originally developed by Chafe (1980), the Pear Film test consisted of a six-minute film and three questions. The film, containing only sound effect but no dialogue, showed the following events: (1) a boy on a bike stealthily took away one basket of pears that were being picked by a man; (2) while the boy rode off, a pretty girl also on a bike approached and passed by him from the opposite direction; (3) as he turned to look at her, his hat flew away, the front wheel of his bike hit a rock; (4) fortunately, three other boys came and helped him pick up the scattered pears; (5) feeling grateful, the boy shared three pears with them. After watching the film, the participants were instructed to orally answer the following questions in five to ten minutes. The questions, provided in both written and spoken forms all at once, were:

1. After watching the Pear Film, please tell me what happened in the film?

2. Please explain what you consider to be the meaning behind the film.

3. Please relate an aspect of the film (any aspect) to something in your own life or something you have observed in real life.

The other speaking test used in the present study was the IELTS sample speaking test, consisting of three parts: (1) general topic task, (2) description task, and (3) discussion task. In part one, the examiner asked the candidates to freely talk about a general topic, such as cold weather. In part two, the participants had to describe any kind of competition that they had ever taken part in. Prior to completing the description task, they were allowed to have one minute to think about what they were going to say. They were also welcome to write down some notes. In the last part, the discussion task, the participants were asked to share some opinions about competitions in school and sporting competitions. They were encouraged to speak as much as they could in about 15 minutes.

The participants' oral responses to the three questions were recorded, transcribed, and analyzed. For the purpose of enhancing the ease and reliability of the scoring, IELTS scoring rubric system was used for both tests, so that the raters only had to stick to one scoring rubrics (i.e., IELTS scoring rubrics) throughout the whole rating process. Furthermore, for both tests, four analytic scoring rubrics were used: (1) fluency and coherence, (2) lexical resource, (3) grammatical range and accuracy, and (4) pronunciation. For each of the four scoring criteria, the maximum possible score was nine. That is, for each of the four scoring criteria, obtaining one point indicated no ability to speak English, whereas getting nine points suggested full command of spoken English. For each test, the subtotal rating score given by each judge for each participant was obtained by adding up the scores across the four score criteria. The average of the two judges' subtotal rating scores was each participant's total score for each test, and the average of the total scores across the two tests was each participant's final total score of his/her overall ESP.

Additionally, for the purpose of increasing the inter-rater reliability between the two raters, a training session was held prior to the formal study. Specifically, the two raters were informed and introduced about the IELTS rating scale and other related materials, such as the purposes and formats of the speaking tests. The two raters were then presented with a set of two speaking samples (i.e., one from the Pear Film speaking test and the other from the IELTS speaking sample test) from a

participant randomly selected. They were asked to rate the samples independently. Then a discussion session was followed to ensure that the two raters could have the same understanding about the scoring rubrics. The inter-rater reliability estimates for scores on the Pear Film speaking test and on the IELTS speaking test were .90 and .98, respectively.

Finally, with respect to AULCK and AUIK for each of the two tests, the scoring criteria employed in the LCT and the TOI were also adopted for identifying acceptable collocations and idioms. Based on the recording and transcribing of the participants' ESP on each speaking test, the frequencies of AULCK and AUIK were obtained by counting the number of times that acceptable collocations and idioms were used by each participant divided by the total number of words that each participant uttered and then multiplied by 100.

Procedures and Data Analysis

The present study contained two sessions. In the first session, the participants were required to take the three paper-and-pencil tests individually in 90 minutes, including the PVLT, the LCT, and the TOI. In the second session, the participants took the two speaking tests individually in a quiet room. Each of the two speaking tests lasted for about 15 minutes. The data collected were then analyzed through the Pearson correlation procedure with the Statistical Package for Social Science (SPSS) version 17.0.

Results

Descriptive Statistics of the Five Instruments

Table 2 presents the descriptive statistics of the participants' performance on the three vocabulary knowledge tests and the two speaking tests, including their maximums, minimums, means, and standard deviations. The mean percentage correct score (75%) of the PVS test was the highest among the three vocabulary knowledge tests, followed by the mean percentage correct score (67%) of the LCT. The TOI obtained the lowest mean percentage correct score (38%). As for the two levels of PVLT, the most frequent 2,000 word level obtained the higher percentage (89%) than the 3,000 word level (62%). In other words, the 2,000 word level test was the comparatively easy for the participants, while the 3,000 word level test was

a bit hard for them to answer.

For the purpose of further examining the participants' performance on the five different categories of LCK (i.e., L1: Verb + Noun, L2: Adjective + Noun, L3: Noun + Verb, L4: Adverb + Adjective, and L5: Verb + Adverb), their means and standard deviations are also displayed in Table 2. Among the five types of lexical collocational knowledge, the L1 type obtained the highest mean percentage correct score (78%), followed by the L3 type (73%). On the other hand, the L4 type obtained the lowest mean percentage correct score (56%). That is, the participants had the best performance on the L1 type. On the other hand, their performance on the L4 type was the worst.

In terms of the mean total rating scores for the three ESP indicators, the participants performed the best on the IELTS speaking test (M = 56%), but the worst on the Pear Film speaking test (M = 54%). However, the differences in the mean rating scores for the three ESP indicators were very trivial.

Table 2 *Descriptive Statistics of the Five Instruments* (N = 32)

Tests	Maximum Possible Score	Minimum (%)	Maximum (%)	Mean (%)	*SD*
PVLT	36	21 (58%)	34 (94%)	26.84 (75%)	3.54
PVLT2000	18	13 (72%)	18 (100%)	15.97 (89%)	1.40
PVLT3000	18	6 (33%)	16 (89%)	11.19 (62%)	2.67
LCT	100	44 (44%)	91 (91%)	67.28 (67%)	10.86
LCT-L1: V + N	20	10 (50%)	20 (100%)	15.50 (78%)	2.20
LCT-L2: Adj + N	20	5 (25%)	19 (95%)	12.50 (63%)	3.78
LCT-L3: N + V	20	8 (40%)	18 (90%)	14.69 (73%)	2.71
LCT-L4: Adv + Adj	20	3 (15%)	18 (90%)	11.13 (56%)	3.71
LCT-L5: V + Adv	20	4 (20%)	20 (100%)	13.47 (67%)	3.75
TOI	50	2.5 (5%)	37 (74%)	19.17 (38%)	9.33
Overall ESP	36	15.8 (44%)	26.3 (73%)	19.88 (55%)	2.39
Pear Film speaking	36	15.5 (43%)	25.5 (71%)	19.52 (54%)	2.12
IELTS speaking	36	15.0 (42%)	27.0 (75%)	20.23 (56%)	2.78

Relationships of ESP to PVS, LCK, and IK

As for the relationships of ESP to the three subcomponents of vocabulary knowledge (i.e., PVS, LCK, and IK), the Pearson correlation analyses were conducted, and the correlation coefficients are shown in Table 3. The participants' PVS (r = .57, p < .01) and IK (r = .54, p < .01) were significantly correlated with their ESP. However, no significant correlation (r = .13, p > .05) was obtained between ESP and LCK.

In terms of the inter-correlations among the three subcomponents of vocabulary knowledge, the results in Table 3 indicate that they were moderately associated with one other. Specifically, PVS was shown to have the strongest correlation (r = .60, p < .01) with IK, followed by its correlation with LCK (r = .50, p < .01). LCK was found to have a low to medium size of correlation coefficient (r = .39, p < .05) with IK.

Table 3 *Correlations between Performance on PVS, LCK, IK and Overall ESP* (N = 32)

Variable	1	2	3	4
Overall ESP	1.00			
PVS	.57**	1.00		
LCK	.13	.50**	1.00	
	.54**	.60**	.39*	1.00

Note. *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Relationship between ESP and PVS

The first research question sought to explore the relationship between ESP and PVS. The Pearson correlation analyses were conducted, and the correlation coefficients are presented in Table 4. As seen from Table 4, in terms of overall PVS, significant and moderate correlations were found with all the three ESP indicators. In particular, the Pear Film speaking performance indicator obtained the strongest correlation with overall PVS, even though the size of the correlation was just moderate (r = .58, p < .01). What came next was the correlation between overall ESP indicator and overall PVS (r = .57, p < .01). By comparison, the lowest correlation was obtained between the IELTS and overall PVS (r = .54, p < .01).

Despite the slight difference in the size of correlation coefficients, the results indicated that in general, the bigger PVS the participants acquired, the higher English speaking scores they tended to obtain.

In an attempt to take a close look at the relationship between overall ESP and PVS, the present study also examined the relationships of ESP to PVS at the 2,000 and 3,000 word levels. In terms of the 2,000 word level, a significant relationship (r = .39, p < .05) was found only for the Pear Film speaking performance. As for the 3,000 word level, significant correlation was found for all the three ESP indicators, with the correlation coefficients ranging from .54 to .67.

Table 4 *Correlations between ESP and PVS* (N = 32)

	Overall PVLT	PVLT2000	PVLT3000
Overall ESP	.57**	.33	.62**
Pear Film	.58**	.39*	.67**
IELTS	.54**	.28	.54**

Note. *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Relationship between ESP and LCK

For the purpose of addressing the second research question concerning the relationship between ESP and LCK, the Pearson correlation coefficients were calculated between the participants' scores on the speaking tests and the LCT. Surprisingly, the results in Table 5 reported a statistically non-significant relationship (r = .13, p > .05) between the participants' overall ESP and their total scores on the LCT. Likewise, neither the Pear Film speaking performance (r = .14, p > .05) nor the IELTS speaking performance (r = .12, p > .05) was found to be significantly correlated with LCK.

A further analysis on the relationships of the five types of lexical collocational to the three ESP indicators revealed an interesting finding. For the overall ESP (r = .36, p < .05) and the IELTS speaking performance (r = .37, p < .05), only L2 type obtained significant but low correlations. In other words, there were no significant correlations of the other four types of LCK to the overall ESP (r = -.23 ~ .08, p > .05) and the IELTS speaking performance (r = -.26 ~ .08, p > .05). As for the Pear Film speaking performance, their relationships to the five types of LCK were all

non-significant (r = -.17 ~ .08, p > .05). The results indicated that only the participants' knowledge about the L2 (Adjective +Noun) collocation type was somewhat related to their ESP.

Table 5 *Correlations between ESP and LCK* (N = 32)

	LCT	L1	L2	L3	L4	L5
Overall ESP	.13	.05	.36*	-.23	.08	.06
Pear Film	.14	.06	.33	-.17	.08	.07
IELTS	.12	.03	.37*	-.26	.08	.05

Note. *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Relationship between ESP and AULCK

In terms of the relationship between the ESP and the overall AULCK, the results in Table 6 indicated that no significant correlation was found for any of the three ESP indicators. In spite of the non-significant correlation coefficients obtained in Table 6, the IELTS speaking performance appeared to have slightly stronger correlations (r = .31, p > .05) with overall AULCK than the other two ESP indicators (r = .09 ~ .22, p > .05).

In terms of the five types of LCK actually used, several findings deserve mentioning here. First of all, only L4 type (Adverb + Adjective) was found to be significantly and moderately correlated with all the three ESP indicators. Specifically, the L4 type displayed moderate correlations with overall ESP (r = .71, p < .01), the Pear Film speaking performance (r = .69, p < .01), and the IELTS speaking performance (r = .70, p < .01). As for the relationship of the other four types of LCK actually used, no significant correlations (r = .03 ~ .29, p > .05) were found for all the three ESP indicators.

Table 6 *Correlations between ESP and AULCK* (N = 32)

	Overall AULCK	L1	L2	L3	L4	L5
Overall ESP	.22	.24	.27	.20	.71**	.06
Pear Film	.09	.16	.25	.26	.69**	.09
IELTS	.31	.29	.28	.16	.70**	.03

Note. *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Relationship of ESP to IK and AUIK

For the final research question concerning the relationship between ESP and IK, correlation analyses were also made and the results are presented in Table 7. As can be seen in Table 7, the participants' IK, as measured by the TOI, was all significantly and moderately correlated with the three ESP indicators, with correlation coefficients ranging from .50 to .56.

For the purpose of finding possible relations between ESP and AUIK, correlation coefficients were further computed and are also shown in Table 7. The results indicated that AUIK was not significantly correlated with any of the three ESP indicators, with the correlation coefficients ranging from -.05 ~ .32. However, one unexpected finding was observed from Table 7. That is, compared with AUIK for the Pear Film speaking test, AUIK for the IELTS speaking test appeared to have slightly lower correlations. In fact, some of them were even negative. For example, the AUIK for the IELTS was negatively but weakly correlated (r = -.05, p > .05) with the Pear Films ESP.

Table 7 *Correlations between ESP and TOI, and between ESP and AUIK* (N = 32)

	TOI	AUIK -Overall	AUIK-Pear	AUIK-IELTS
Overall ESP	.54**	.22	.32	-.03
Pear Film	.56**	.20	.32	-.05
IELTS	.50*	.22	.30	-.02

Note. *. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Discussions and Conclusions

One of the research purposes in the current study was to investigate the relationship between overall PVS and ESP. The participants' overall PVS was found to be moderately correlated with their ESP assessed by the three ESP indicators, with the correlation coefficients ranging from .58 to .54. The findings seemed to be consistent with that of Zimmerman's (2004) study, which reported a significant and slightly higher correlation (r = .66, p < .0001) between ESL learners' PVS and their overall ESP. Despite the slight difference in the magnitude of the correlation coefficients between the two studies, the findings of the present study provide

additional evidence to corroborate the longstanding claim that PVS plays a somewhat influential role in ESP.

In terms of the relationships between overall LCK and the three ESP indicators, no significant relationship was found. The somewhat surprising results appeared to contradict both Sung's (2003) and Chiu and Hsu's (2008) results. Both Sung's (r = .78) and Chiu and Hsu's (r = .56) studies showed significant and moderate correlations between overall ESP and overall LCK. A plausible explanation for these contradictory findings might be due to the differences in sample sizes and the measures used to score ESP. The present study recruited only 32 students, whereas the numbers of participants were 72 and 56, respectively, in Sung's and in Chiu and Hsu's study. According to Kiess (2002), correlation coefficients are very sensitive to sample size. Thus, since the two previous studies (Chiu & Hsu, 2008; Sung, 2003) included more participants, it was easier to obtain a significant correlation between ESP and LCK. Moreover, the measures used in the present study to score ESP were not exactly the same as those in the two studies. Specifically, the present study used the Pear Film speaking test and the IELTS Speaking subtest. In Chiu and Hsu's study, in addition to the Pear Film test, the PhonePass spoken English test was also used, which was scored automatically by a computer-based system with scale points ranging from 20 to 80 and (as mentioned earlier) had problems with its construct validity. As to Sung's study, the rating scale of the Pear Film test ranged from 20 to 60 scale points. However, the present study adopted a zero-nine analytic scoring scale for the two speaking tests. As asserted by Bachman and Palmer (1996), different characteristics of language tasks, such as scoring methods or item formats, could affect how the scores of the test reflect the targeted construct (i.e., ESP). Given the fact that scoring methods used in the present study were different from those of the previous studies, differences existing in the results among the studies would not be too surprising. Thus, the relationship between ESP and overall LCK seemed to remain inconclusive at this point. It awaits further studies to incorporate multiple measures for verify the relationship between learners' ESP and LCK.

A close look at the relationships between the five types of LCK and the three ESP indicators revealed an interesting finding. That is, the L2 (Adjective + Noun) was the only type found to have significant correlations with the overall ESP indicator (r = .36, p < .05) and the IELTS speaking performance indicator (r = .37, p < .05), despite the small size of correlation coefficients. The findings appeared to be

inconsistent with those of Sung's (2003) and Chiu and Hsu's (2008) research. In Sung's (r = .50 ~ .72) and Chiu and Hsu's (r = .31 ~ .44) studies, in addition to the L2 type, significant correlations were also found with various types of LCK. As mentioned earlier, the possible reason for the inconsistent findings might be the different sample sizes. At this point, this inconsistency awaits further studies to verify. Based on the finding of the present study, it appeared that learners' performance on the Adjective + Noun type (i.e., the L2 type) of LCK appeared to be somewhat related to their ESP. The reason, however, awaits further studies to explore.

Similar to LCK, nor was AULCK found to have any significant correlation with the three ESP indicators. The finding seemed to be in line with that of Chiu and Hsu's (2008) study, where no significant correlation was found between ESP and AULCK. However, the findings of the two studies appeared to be contradictory to those of other studies (e.g., Hsu, 2007; Sung, 2003; Ushigusa, 2008), where significant and moderate correlations were found between ESP and AULCK. One possible explanation for the inconsistent results among the studies might be due to the difference in the number of collocational knowledge types involved. For instance, Ushigusa (2008) adopted the classification of collocations, which consisted of both grammatical and lexical collocations. As Shin and Nation (2008) noted from a corpus study, grammatical collocations were found to be the most frequent collocations of English. That is, grammatical collocations tend to be used very often and thus may account for the most proportion of collocations in one's speech. As such, the frequency of AULK would tend to become higher and more varied in Ushigusa's study, which would in turn increase the correlation coefficient (r = .46, p < .05), between the actual use of collocations and ESP (Kiess, 2002). By contrast, the present study included only lexical collocations, which would result in lower frequency of collocations actually used, smaller variance, and smaller correlation coefficient. Therefore, the relationship between ESP and AULK remains inconclusive. Future studies are warranted to further probe into the relationship between the two variables.

One more interesting finding deserves some discussion. That is, significant and moderate relationships of AULK in L4 (Adverb + Adjective) were found with each of the three ESP indicators, with the correlation coefficients ranging from .69 to .71. A plausible explanation for the findings might be due to the fact that the present study included the IELTS speaking test, where the participants had to tell their personal experience about competitions. They tended to use lots of adjectives

to portray how they felt in that event, and those words used were counted as L4 type of collocation. For instance, "very interesting" and "pretty good" were the common phrases used by the participants in the two speaking tests.

With regard to the relationship between IK and ESP, significant and moderate relationships (ranging from .50 to .56) were obtained in the present study for all the three ESP indicators, which provided additional supporting evidence for the presumed relationship between IK and ESP. As claimed by Nippold and Martin (1989), possessing IK is vital for learners in oral communication.

As to AUIK, non-significant and weak correlation coefficients were obtained in the present study with each of the three ESP indicators, with the correlation coefficients ranging from -.05 ~ .32. The results of the present study appeared to be in agreement with that of Chiu and Hsu's (2008) study, where an insignificant correlation (r = .25, p > .05) was found between the learners' AUIK and their ESP. However, a significant and moderate correlation (r = .53, p < .05) between AUIK and ESP was found in Ushigusa's (2008) study. A plausible explanation for the contradictory findings among the studies might be due to the difference in defining IK. In the present study, IK was defined as fixed, non-literal idioms, whereas in Ushigusa's (2008) study, it consisted of two sub-categories, phrasal verbs and idiomatic multiword units, which tended to be used quite often in speaking (Wray, 2002). Therefore, as more categories of idioms were included in Ushigusa's study, the frequency of idioms actually used would become higher and more varied, and so did the correlation coefficient (r = .53, p < .05) between the frequency of idioms actually used and ESP (Kiess, 2002). Due to the inconsistency in the results between the two studies, it remained inconclusive about the relationship between ESP and AUIK. Therefore, a need is in order for future studies to delve further into the relationship between the two variables.

Conclusions

Based on the results of the present study, the following conclusions are made as follows in response to the three research questions. First of all, there was a significant but moderate correlation between the overall PVS and ESP. Next, the overall LCK was not significantly correlated with ESP. Likewise, no significant relationship was found between AULCK and ESP. Finally, IK was moderately related to ESP. However, the relationship of AUIK to ESP was not strong enough to

be significant.

Pedagogical Implications

Based on the results of the present study, several pedagogical implications can be made for English speaking instruction. First of all, as the current study shows, PVS is somewhat related to ESP, suggesting that teachers could integrate vocabulary instruction in their teaching syllabi of speaking courses. Moreover, as Webb (2008) indicated, knowing students' PVS enables teachers to determine whether the students can speak confidently and proficiently about a certain topic. Before implementing speaking activities, teachers should ensure that students have enough vocabulary knowledge. After learners have enough vocabulary, teachers can further help them to put the vocabulary items to productive use through a range of activities. For example, if the learners have difficulty in recalling or saying a word that they have already learned, the teachers can provide a context for the word and ask the learners to guess the word.

Though the present study failed to find the significant relationship between overall LCK and ESP, the LCK of the L2 (Adjective + Noun) type was found to have significant correlations with the ESP indicators. Significant correlations were also obtained between the actual use of the L4 (Adverb + Adjective) type and the ESP indicators. Thus, it is still worthwhile to address the importance of LCK when EFL teachers conduct speaking instruction. Besides enhancing students' vocabulary size, language teachers should also provide explicit instruction on students' LCK. Typical awareness-raising techniques are recommended by Lewis (1997), such as asking students to keep a notebook of lexical collocations, and getting students to pick out the words which do not strongly collocate with the targeted word from a list of words. Moreover, language teachers can ask students to think up as many collocations of *Adjective* + *Noun* and *Adverb* + *Adjective* as they can with a common word (e.g., money: promising/ great/ bright future; wrong: totally/ completely/ absolutely/ obviously wrong). Teachers can also use textbooks and other supplemental materials to train students to identify lexical collocations of *Adjective* + *Noun* and *Adverb* + *Adjective* and then lead students to do some follow-up practice (e.g., matching games and collocation grid exercises). By doing so, students may acquire LCK and ultimately use it while speaking.

Additionally, based on the significant but moderate correlations found in the present study between IK and ESP indicators, it seems that idiom instruction plays

an essential role in speaking syllabi. As suggested by Nation (2001), when teaching an idiom, teachers should first introduce an idiom, explain its underlying meanings, and then provide its relevant cultural clues, rather than just offer the L1 equivalent translation. After students can recognize the idioms, teachers should expose students to the idioms in the authentic materials, such as movies, commercials and English songs, so that students can easily recall the actual meaning of the idioms in a certain context. Another worth-trying strategy, as advocated by Hulstjin (1997), is the keyword method. Specifically, when teaching unfamiliar idioms, teachers can use this memory technique to help students make a link between the targeted new idioms and either some known words or an image. The link can facilitate the students to retrieve the targeted idioms. As idioms often refer to some concrete event or phenomenon, the keyword method is assumed to be suitable for teaching IK.

Limitations of the Study

The generalizability of the present study's results is limited by the small number and the nature of the participants. As the present study involved assessment of not only 90-minute paper-and-pencil tests but also 30-minute oral tests, recruiting a large number of the participant appeared to rather difficult. It turned out only 32 participants were involved in the present study. Moreover, the participants were the graduate students majoring in English. Therefore, it remains unknown about whether the results can be generalized to a larger number of students with different English proficiency levels or different majors. In other words, the results of the current study only could be generalized to graduate students who major in English and have a similar English proficiency level.

In addition, the generalizability of the present study was also subject to the speaking instruments. The present study adopted two speaking tests (e.g., a film description task and a description/discussion task about personal experience in participating in any competitions) to assess the participants' overall English speaking performance. However, different speech topics and various types of speaking tasks would lead to the learners' different speaking performance (Skehan & Foster, 1999). Thus, the results of the speaking tests might not be fully representative of the participants' overall speaking ability.

Another concern also about measurement is that the scores of the LCT might not fully reflect LCK. The present study only included the five types of lexical

collocations (i.e., Verb + Noun, Adjective + Noun, Noun + Verb, Adverb + Adjective, and Verb + Adverb), which differed from the seven types of lexical collocations originally proposed by Benson et al. (1997). In other words, the participants' knowledge and use of grammatical collocations were not assessed in the present study. As mentioned before, grammatical collocations account for a large proportion of speech. Hence, it awaits future studies to investigate whether grammatical collocational knowledge plays an important role in successful speaking performance. Hence, it is advised that future investigations could include grammatical collocations as well in their studies.

One more measurement-related limitation is that the TOI employed in the present study may fail to fully reflect learner's IK. As mentioned earlier, the idioms selected were based on the two criteria: (1) the idioms must be found from various senior high school English textbooks commonly used in Taiwan, and (2) the words in the idioms selected were limited to the MOE 1,000 Basic English words list. Other idioms failing to meet the two criteria were bound to be excluded, and thus could not be tested although the participants may answer them right.

Suggestions for Future Research

While this study had its limitations, it was hoped that it could offer a basis for future research on the relations of ESP to the three components of vocabulary knowledge. Some directions for further studies are recommended in the following.

First of all, the study could be replicated to investigate whether the results of the present study could be confirmed with a larger number of participants of various majors and different proficiency levels. Second, when encountering different topics or speaking tasks, such as picture description and story telling, learners' speaking performance could differ. Thus, it needs to be further investigated whether the results can also be applicable to studies using other speaking tasks or tests. Moreover, future research could entail a wider variety of scoring rubrics. As suggested by Higgs and Clifford (1982), different scoring measures may lead to different scores for the same level of overall language proficiency. Thus, it is recommended that future studies could adopt multiple systematic scoring methods to fully examine and gauge the various aspects of English speaking ability.

As for the measurement of vocabulary knowledge, the current study investigated only the three subcomponents of vocabulary knowledge. As recommended by Hsu (2007), to obtain a complete understanding about relationship

between vocabulary knowledge and ESP, future studies along the line should incorporate other vocabulary knowledge subcomponents, such as knowledge of word association, derivational morphological knowledge. Furthermore, since the present study was restricted to lexical collocational knowledge, it was also suggested that future studies could include grammatical collocational knowledge in order to fully understand the relationship between EFL learners' ESP and collocational knowledge. That is, given a weak correlation obtained in the present study between lexical collocational knowledge and ESP, a further step could be taken to re-examine collocational knowledge with a bigger scope.

In addition to vocabulary knowledge, for the purpose of getting a clear picture of the English speaking construct, future research could explore its relationship to other variables, such as speaking task types (Foster & Skehan, 1996), planning time (Ortega, 1999), and affective reactions to speaking tests (Scott, 1986). It may also be important to see whether task characteristics or speaker's characteristics have effects on the nature of ESP (Teng, 2007). By incorporating these variables, the results of future studies could help extend the understanding about ESP.

References

Aghbar, A. A., & Tang, H. (1991, March). *Partial credit scoring of cloze type items*. Paper presented at 1991 Language Testing Research Colloquium, Education Testing Service, Princeton, NJ.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In. J. Guthrie (Ed.), *Comprehension and teaching research reviews* (pp. 77-117). Newark, ED: International Reading Association.

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

Benson, M., Benson, E., & Ilson, R. (Eds.). (1997). *The BBI dictionary of English word combinations.* Amsterdam: John Benjamins.

Berman, R., & Cheng, L. (2000). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics*, *4*, 25-40.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finnegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.

Chafe, W. L. (1980). The deployment of consciousness in the production of a narrative. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 9-50). Norwood, NJ: Ablex.

Cheng, C. K. (2007). College English majors' receptive and productive vocabulary knowledge. In English Teachers' Association of the Republic of China (Ed.), *The proceedings of Sixteenth International Symposium and Book Fair on English Teaching* (pp. 347-356). Taipei: Crane.

Chiu, C. Y., & Hsu, J. Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *The Asian EFL Journal*, *10*, 181-204.

Chuang, Y. Y. (2010). A study of Taiwanese technical college students' affective reactions to speaking English in the EFL classroom. *Journal of National Huwei University of Science and Technology*, *29*(1), 75-94.

Ellis, N. C. (1997). Vocabulary acquisition: Word structure, collocation, word-class. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 122-139). Cambridge: Cambridge University Press.

Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instructions* (pp. 33-68). Cambridge: Cambridge University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*, 299-323.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing collocations and formulae. In A. P. Cowie (Ed.), *Phraseology, theory, analysis, and applications* (pp. 145-160). Oxford: Clarendon Press.

Higgs, T., & Clifford, R. (1982). The push towards communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57-59). Lincolnwood, IL: National Textbook Company.

Hsu, B. T. (2009). *The correlation between senior high school students' vocabulary*

size and reading performance (Unpublished master's thesis). National Taiwan Normal University, Taiwan.

Hsu, J. Y. (2006). An analysis of the multiword lexical units in contemporary ELT textbooks. In Department of English Wenzao Ursuline College of Languages (Ed.), *The proceedings of the 23rd International Conference on English Teaching and Learning in the Republic of China* (pp. 363-381). Kaohsiung: Kaun Tang International Publications Ltd.

Hsu, J. Y. (2007). *Multiword lexical units and their relationship to impromptu speech*. Retrieved from ERIC database. (ED495303)

Hulstjin, J. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203-224). New York, NY: Cambridge University Press.

IELTS 7. (2009). Cambridge: Cambridge University Press.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, *29*, 24-49.

John-Laird, P. N. (1993). Foreword. In C. Cristiana & R. Patricia (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 1-5). New Jersey: Lawrence Erlbaum Associates.

Kiess, H. O. (2002). *Statistical concept for the behavioral sciences* (3rd ed.). Boston, MA: Allyn and Bacon.

Koizumi, R. (2005). *Relationships between productive vocabulary knowledge and speaking performance of Japanese learners of English at the novice level* (Unpublished doctoral dissertation). University of Tsukuba, Japan.

Laufer, B., & Nation, P. (1999). A productive-size test of controlled productive ability. *Language Testing*, *16*, 33-51.

Lea, D. (Ed.). (2002). *Oxford collocations dictionary for learners of English*. Oxford: Oxford University Press.

Lewis, M. (1997). *Implementing the lexical approach*. Hove: Language Teaching Publications.

Liu, H. Y. (2006). *Relationship between EFL learners' English proficiency levels and three types of word knowledge* (Unpublished master's thesis). National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan.

Liu, Y. F. (2002). *The effect of English word exposure frequency in the simplified novel on incidental vocabulary acquisition by EFL senior high school students in Taiwan* (Unpublished master's thesis). National Kaohsiung Normal University, Kaohsiung, Taiwan.

McCarthy, M. (2004). *From corpus to course book*. Cambridge: Cambridge University Press.

McCatig, I. R. (Ed.). (1994). *Oxford learner's dictionary of English idioms*. Oxford: Oxford University Press.

McDevitt, E. (1993). *What does that mean? An introduction to American idioms.* Retrieved from ERIC database. (ED372660)

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nippold, M. A., & Martin, S. T. (1989). Idiom interpretation in isolation versus context: A developmental study with adolescents. *Journal of Speech and Hearing Research*, *32*, 59-66.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, *20*, 109-148.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 209-227). Amsterdam, Netherlands: John Benjamins.

Scott, M. L. (1986). Student affective reactions to oral language tests. *Language Testing*, *3*, 99-118.

Shin, D. (2007). The high frequency collocations of spoken and written English. *English Teaching*, *62*, 199-218.

Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, *64*, 339-348.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*, 93-120.

Sung, J. (2003). *English lexical collocations and their relation to spoken fluency of adult non-native speakers* (Unpublished doctoral dissertation). Indiana University of Pennsylvania, Pennsylvania.

Teng, H. C. (2007). *A study of task types for L2 speaking assessment*. Retrieved from ERIC database. (ED496075)

The Language Training & Testing Center. (2012). *Score data summary for 2012 GEPT*. Retrieved from http://www.lttc.ntu.edu.tw/results.htm

Tsai, B. Y. (2005). *The relationship between receptive English vocabulary sizes and listening comprehension competence of college EFL students* (Unpublished master's thesis). National Kaohsiung Normal University, Kaohsiung, Taiwan.

Ushigusa, P. (2008). *The relationships between oral fluency, multiword units, and proficiency scores* (Unpublished doctoral dissertation). Purdue University, Indiana.

Webb, S. (2008). Receptive and productive sizes of L2 learners. *Studies in Second Language Acquisition*, *30*, 79-95.

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

Wray, A. (2002). *Formulaic language and lexicon*. Cambridge: Cambridge University Press.

Zimmerman, K. J. (2004). *The role of vocabulary size in assessing second language proficiency* (Unpublished master's thesis). Brigham Young University, Provo, Utah.

英語口說表現與字彙量、搭配詞知識及成語知識之關係研究

曾思菱* 林文鶯**

本研究旨在探討臺灣英語學習者的口說表現與字彙量、搭配詞、成語知識的關係。本研究主要回答以下三個研究問題：一、字彙量知識與口說表現的相關程度為何？二、搭配詞知識與口說表現的相關程度為何？三、成語知識與口說表現的相關程度為何？研究對象為就讀或畢業於北臺灣英語相關系所的 32 位碩士生。研究工具包含五項測驗：一、字彙量測驗；二、搭配詞測驗；三、成語測驗；四、影片口語測驗；五、雅思口語測驗。資料分析的工具則採用基本敘述統計、皮爾森相關係數。研究結果顯示英語所碩士生的口說表現與字彙量知識及成語知識的相關性為顯著正相關。但搭配詞知識與口說表現之間卻沒有顯著關聯。另外受試者使用搭配詞知識和成語知識也沒有和口說表現呈顯著關聯。針對於本研究的結果，本研究最後提供對英語教學之應用，以及對未來研究的建議。

關鍵詞：口說表現、字彙量、搭配詞、成語知識

* 作者現職：新北市頭湖國小英語老師

** 作者現職：臺北市立大學英語教學系副教授

通訊作者：林文鶯，e-mail: wylin@utaipei.edu.tw