Lexical Collocations and Their Impact on the Online Writing of Taiwanese College English Majors and Non-English Majors

Jeng-yih (Tim) Hsu
Department of English
National Kaohsiung First University of Science and Technology
gogotim@ccms.nkfust.edu.tw

Abstract

The present study investigates the use of English lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. Data for the study were collected from 41 English majors and 21 non-English majors at a national university of science and technology in southern Taiwan. Each student was asked to take a 45-minute online English writing test, administered by the web-based writing program, Criterion Version 7.1 (Educational Testing Services) to examine the subjects’ use of lexical collocations (i.e., frequency and variety). The test was also used to measure writing fluency of the two student groups.

Test results were examined to answer the two major questions for correlation (1) between the subjects’ frequency of lexical collocations and their writing and (2) between the subjects’ variety of lexical collocations and their writing. The study findings indicated that: (1) there seemed to be a positive correlation between Taiwanese college EFL learners’ frequency of lexical collocations and their online writing scores; and (2) there seemed to be a significantly positive correlation between the subjects’ variety of lexical collocations and their online writing scores. This present study also reports on a pattern of lexical collocation development observed among writers of different fluency levels, ranging from the lowest to the highest.

Keywords: lexical collocations, on-line writing, ETS Criterion, EFL students
1. Introduction

1.1 Status of English Collocations

It has been fourteen years since the appearance of Lewis’ (1993) book, The Lexical Approach, directing our attention to systematically reexamine the importance of collocations in EFL/ESL education. Along with Lewis, scholars of second language vocabulary acquisition, ESL/EFL material and curriculum design, pedagogy, and lexicography, such as Bahns and Eldaw (1993), Benson, Benson, and Ilson (1997), Ellis (1996, 2001), Howarth (1998), McCarthy (2004a, 2004b), Nation (2001), and Richards and Rodgers (2001), have also joined the same trend of arguments, advocating the necessity to study English collocations.

First brought up by Palmer (1933) and later introduced to the field of theoretical linguistics by Firth (1957), the definition of collocations most commonly shared is: the tendency of one word to co-occur with one or more other words in a particular domain (Aghbar, 1990; Al-Zahrani, 1998; Gitsaki, 1999; Nation, 2001; Nesselhauf, 2003). In his monograph Second Interim Report on English Collocations, the father of collocation studies, linguist Palmer (1933) simply puts, “Each [collocation] … must or should be learnt, or is best or most conveniently learnt as an integral whole or independent entity, rather than by the process of piecing together their component parts” (Palmer, 1933, p. 4, cited in Nation, 2001, p. 317). Collocation itself is far from being thoroughly described or understood whether we approach it from a semantic, syntactic, or idiomatic approach (Hsu, 2002).

We all know collocation is crucial and recognize it is the collocational proficiency that differentiates native and non-native speakers (Ellis, 2001; Koya, 2006; McCarthy, 2004a; Nation, 2001; Wouden, 1997) but what we can do as classroom teachers is not much. Commercial coursebook writers, echoing Lewis’ idea (Lewis, 1993; 1997; 2000) have down poured a wide collection of new materials, all of which are ‘promoting’ the teaching and learning of collocations. Good examples include ESP textbooks—Meeting (Goodale, 1993), Business Matters (Powell, 1996a), Presenting in English (Powell, 1996b), Build Your Business Grammar (Bowen, 1997), and The Working Week (Watson-Delestree & Hill, 1998) and EGP (i.e., English for General Purposes) materials of recent years—Communication Strategies (Paul, 2003a) and Further Communication Strategies (Paul, 2003b), Touchstone: Book I, II, III, & IV (McCarthy, McCarten, & Sandiford, 2005), and Totally True: Book I, II, & III (Huizenga & Huizenga, 2005). When turning for help from researchers who conducted series of empirical studies, their reports on English collocations are too diverse, yielding findings of limited practical values. For instance, many pioneer studies merely investigated EFL/ESL learners’ knowledge of collocations (Bahns & Eldaw, 1993; Farghal & Obiedat, 1995; Zughoul & Abdul-Fattah, 2003) or analyzed how EFL/ESL learners committed errors1 of collocation (Huang, 2001; Jeng, 2006; Liu, 1999a; 1999b; Lombard, 1997; Wu, 2005). Only until recent have we seen studies targeting the possible connection between collocations and general language skills (Al-Zahrani, 1998; Bonk, 2000; Sung, 2003).

Nevertheless, with the newly invented technology becoming available, the field of EFL/ESL is seeing many revolutionary changes. Among many, computer-assisted language learning has been the issue in the core discussion. To be more specific, the use

---

1 For a detailed list of empirical studies on collocation error analysis, see Appendix A.
of internet in EFL/ESL classrooms is pushing both teachers and students to a new generation of challenges. The adaption of internet- or web-based materials is spreading out into not only the teaching and learning but also assessment aspects. Our second language learners are now being both taught and tested by online materials. A decade ago when we were still in paper-and-pencil period, the very first empirical study on collocations, conducted by Zhang (1993), fell into the area of EFL/ESL writing. The followers of Zhang took several steps further, investigating the relations between collocations and general English proficiency (Al-Zahrani, 1998; Bonk, 2000), between collocations and speaking (Sung, 2003), between collocations and reading (Lien, 2003), and between collocations and listening (Hsu & Hsu, 2007). It is about the right time to start looking into how the acquisition of collocations may or may not affect each of EFL/ESL learners’ language skills, including writing, speaking, reading, and listening in the new era of computer-mediated communication (i.e., CMC).

1.2 Significance of the Current Study

The current study specifically investigates the use of English lexical collocations and their relation to the online writing of Taiwanese college students (i.e., English and non-English majors). The underlying assumptions behind this study design are three-fold. First, by focusing on writing, this study, like Zhang’s study (1993), intends to lead off a series of empirical studies that will examine the connections between collocations and the other language skills in a CMC generation. In addition, the development of computers, internet, and many e-learning materials have shaped EFL/ESL education dramatically. Learning to use and interact in English ‘computer-mediated appropriately’ is becoming increasingly important. Furthermore, the pressure for test takers to conduct timed, computer- or internet-based ‘speed writing’ has never ceased. New English proficiency exams, including TOEFL-iBT, TOEIC2, are now asking students to take an instant writing test online.

It is therefore the hope of this study (and certainly many EFL/ESL writing teachers) to know the role collocations play in our student writers’ online writing, which is being severely judged and evaluated by the latest web-based writing programs3. This study intends to investigate if there is a correlation between Taiwanese EFL college students’ use of lexical collocations (i.e., frequency and variety) and their online writing tested by a web-based writing program (in this case, Criterion 7.1). Test results are later examined to answer the two major questions for correlation (1) between the subjects’ frequency of lexical collocations and their writing and (2) between the subjects’ variety of lexical collocations and their writing.

2. Literature Review

2.1 Historical Background of Collocation Studies

Discussions on collocations began as early as in the 1980s (Celce-Murcia, 2001). However, the initial reports were mostly made by teachers-in-practice who occasionally suggested that it was urgent to implement the teaching of collocations in EFL/ESL

---

2 Educational Testing Services (henceforth, ETS) has announced that a new computer-based writing test component is scheduled to be implemented in 2008.

3 The term, web-based writing program, is first used in Cheng (2006).
classrooms. For example, Smith (1983) and Channel (1981) both asserted that vocabulary instruction ought to include collocations. They believed that collocations provide ‘contexts’ and ‘chances for word associations,’ both beneficial for learning new words. Starting from the 1990s, series of discussions on the essentiality of collocations came to the attention of second language studies and EFL/ESL education because of four main reasons. First of all, scholars of TESOL methodology constantly criticized the fact that the major foreign language teaching methods or approaches (e.g., Grammar Translation Method and Communicative Approach) downplay the role of vocabulary. We needed efficient ways to present vocabulary to learners of foreign languages (Zimmerman, 1997; Schmitt, 2000). Second, the British applied linguist, Lewis initiated his talks on adapting a collocation-based syllabus (1993, 1997) as well as the Lexical Approach (1993). In addition, the availability of online corpora, such as British National Corpus and American National Corpus, guaranteed immediate access to a large collection of authentic written/spoken English examples. Lastly, the compiling of the *BBI dictionary of English word combinations* (Benson, et al., 1997) and *Oxford Collocations Dictionary for Learners of English* (Lea, et al., 2002) helped EFL/ESL practitioners and learners obtain handy references for English collocations.

### 2.2 Empirical Studies on the Relation between Collocations and English Fluency

Empirical studies on collocations did not start until the first doctoral dissertation (Zhang, 1993) came along. Under the advising of Aghbar4 (personal communication, 2002), Indiana University of Pennsylvania, Pennsylania, USA, Zhang set off a trend of collocation studies in which the correlations among the EFL students’ knowledge of collocations (i.e., measured by a fill-in-the-blank test), use of collocations (i.e., frequency of accurate collocations recorded in student English), and English fluency were examined. The similar research design was later modeled by Al-Zahrani (1998) and Sung (2003).

In Zhang’s case (1993), he tested 60 (i.e., 30 natives and 30 non-native) speakers of English at Indiana University of Pennsylvania by using one blank-filling collocation test and one paper-based writing test. The collocation test was measuring the students’ collocational knowledge while the writing task was collecting their use of collocations and writing fluency. In this study, Zhang found that (1) native writers performed significantly better than non-native writers in the collocation test and (2) native writers outperformed the non-native writers in terms of the use of collocations in writing. Zhang concluded: (1) “collocational knowledge is a source of fluency in written communication among college freshmen, and (2) quality of collocations in terms of variety and accuracy is indicative of the quality of college freshmen writing” (p. v).

Following Zhang, Al-Zahrani (1998) studied the knowledge of English lexical collocations among four academic levels of 81 Saudi EFL students and the relationship between the knowledge of lexical collocations and the participants’ general language proficiency. In the study, an English collocation test consisting of 50 verb + noun lexical collocations, a writing test and an Institutional Version of paper-and-pencil TOEFL test were administered. Al-Zahrani’s findings included: (1) a significant difference existed in his students’ knowledge of lexical collocations among the different academic years. To be more specific, the knowledge of lexical collocations increased along with the subjects’

---

4 Professor A. Aghbar is also the Ph.D. dissertation advisor to the author of this paper.
academic years. (2) A strong correlation was found between the students’ knowledge of collocations and their overall language proficiency as measured by the TOEFL test. (3) The writing test was a slightly better predictor of the students’ knowledge of lexical collocations than the TOEFL test.

Al-Zahrani’s study differed from Zhang’s in that the use/frequency of collocations was overlooked; therefore, no report was made in Al-Zahrani’s regarding his subject’s frequency of collocations which could have been observed in the writing test. Later, moving away from EFL learners’ writing skill, Sung (2003) directed our focus to the possible connection between collocations and speaking fluency. She explored EFL students’ knowledge and use of English lexical collocations in their relation to speaking fluency. In her experiment, 24 native and 72 non-native speakers took two tests: one collocation test and one speaking test. The collocation test was assessing the subjects’ knowledge of lexical collocations while the speaking test was eliciting the subjects’ use of lexical collocations and measuring their speaking fluency. Her findings indicated that a significantly strong correlation existed between the EFL students’ knowledge of lexical collocations and their speaking fluency. Sung further concluded: “…knowledge of lexical collocations is a more significant indicator of degree of speaking of proficiency…”(p. v).

In summary, this series of correlation-oriented empirical studies stopped after only the initial exploration on the relations between EFL learners’ collocational knowledge and use, and their writing skill (Zhang, 1993), speaking skill (Sung, 2003), and overall English skill (Al-Zahrani, 1998). Other individual language skills, such as listening, reading, or vocabulary, have not been examined. A list of these studies is summarized in Table 1 for a clear comparison.

<table>
<thead>
<tr>
<th>Researchers (Year)</th>
<th>Subjects</th>
<th>Instruments</th>
<th>Focuses</th>
</tr>
</thead>
</table>
| Zhang (1993)      | 1. 30 natives  
2. 30 non-natives at a university of U.S.A. | 1. collocation test (fill-in-the-blank)  
2. writing task | 1. lexical collocations  
2. grammatical collocations |
2. writing task  
3. TOEFL test | verb + noun lexical collocations |
| Sung (2003)       | 1. 24 natives  
2. 72 non-natives at a university of U.S.A. | 1. collocation test (fill-in-the-blank)  
2. one speaking test | lexical collocations |

### 2.3 Empirical Studies on the Effect of Direct Collocation Instruction in College/University Settings of Taiwan

In the EFL community of Taiwan, interestingly, only four empirical studies related to collocations have been carried out by teacher-researchers in local college/university settings. More interestingly, the consensus of these four studies was reflected by their emphasis on the explicit instruction of collocations in real-time classroom practices. Liu (2000) was the first in Taiwan who studied the impact of collocation instruction on her college EFL students’ writing. In an eighteen-week semester, 49 freshmen English majors
were asked to write a paper-based essay in class at the beginning and the end of the semester. The two essays were later analyzed and compared for the patterns of accurate and inaccurate lexical collocations. Liu’s findings showed that her students produced a greater number and variety of accurate lexical collocations in their second essays. In terms of the causes of the inaccurate lexical collocations in both essays, according to Liu, resulted form the mother tongue interference.

In a private university of science and technology, Hsu (2002) set up an intensive business English course to investigate whether direct collocation instruction would enhance Taiwanese college EFL learners’ development on the knowledge and use of collocations, as well as general language fluency. Altogether 7 English and 2 business majors participated in this 4-week intensive summer course. As a teacher-researcher, Hsu analyzed students’ writings, pre- and post-course collocation tests, collocation notebooks, teacher’s notes, student interviews, and videotapes of the course sessions. Drawing from both quantitative and qualitative data, Hsu found that explicit emphasis on collocations seemed to have helped students acquire new collocations in their written and spoken English. Furthermore, he found that a strongly positive correlation existed between the students’ quantity of lexical collocations and their language fluency observed at a certain time and that a slightly positive correlation also existed between the students’ gain in the quantity of lexical collocations and their gain in the language fluency upon the end of the course. Based on the findings, he recommended that collocation instruction was highly applicable and should be incorporated into every EFL classroom.

Also conducting a study in a Taiwanese college classroom, Lien (2003) addressed the effects of collocation instruction on reading comprehension. A total of 85 Taiwanese college English majors from three academic levels (i.e., sophomores, juniors, and seniors) participated in the study. In the pre-experimental stage, all of the subjects took one collocational test and reading comprehension test. In the experimental stage, the subjects received three types of instruction (i.e., collocation instruction, vocabulary instruction, and no instruction) and took three reading comprehension tests, which were used to measure the effectiveness of the three types of instruction. In the post-experimental stage, a questionnaire was administered to collect the subjects’ attitudes toward the three types of instruction. Lien’s showed that collocation instruction helped her subjects’ reading comprehension progress more evidently than the other two types of instruction did. The majority of the 85 subjects also reported positive attitudes toward the direct teaching of lexical collocations.

Working with a group of 46 sophomores of English majors in a private 5-year junior college, Chu (2006) explored her students’ progress in acquiring collocations based on the two treatments—the noticing of collocations (as presented by the output activities) and the learning of collocations (as presented by the output-input activities). In a one-week period, she intended to investigate if the two types of instructional activities could help her students’ acquisition of new lexical collocations. After comparing the results from the pre- and post-course tests, Chu was convinced by her students’ post-treatment performance especially in the group who received the output activities. She went on to conclude that ‘noticing’ played an important role in the process of acquiring second language (i.e., English, in her case) collocations.

In short, empirical studies researching the impact of collocation instruction on Taiwanese EFL learners have been conducted in several aspects of language skills,
including writing (Liu, 2000; Hsu, 2002), vocabulary development (Hsu, 2002; Chu, 2006), reading (Lien, 2003), overall language fluency (Hsu, 2002). Among these studies, most of the results showed that direct collocation instruction somehow helped improve Taiwanese college EFL learners’ language performance. Specifically, the subjects from the studies maintained a positive attitude toward collocation instruction. The teaching-oriented empirical studies of collocations concerning Taiwanese EFL learners of college/university settings are summarized in Table 2.

Table 2
Empirical Studies on the Effect of Direct Collocation Instruction in College/University Settings of Taiwan

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Subjects</th>
<th>Instruments</th>
<th>Focuses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liu (2000)</td>
<td>49 university students</td>
<td>pre- and post-course essay writing</td>
<td>lexical collocations</td>
</tr>
</tbody>
</table>
| Hsu (2002)  | 9 university students | 1. written assignments  
2. pre- and post-course tests: fill-in-the-blank collocation tests, writing and speaking tests  
3. students’ collocation notebook  
4. teacher class notes  
5. student interviews  
6. videotapes | lexical collocations |
2. reading tests  
3. questionnaire | lexical collocations |
| Chu (2006)  | 46 five-year junior college students | 1. output activities  
2. output + input activities | 1. lexical collocations  
2. noticing and learning of collocations |

3. Methodology

3.1 Research Design

In this study, the data were collected by one test—Criterion (7.1 Version) web-based writing program. Three variables elicited from the same test were used. The use of lexical collocations was first recorded based on the raw count of lexical collocations from the subjects’ writing test. This raw quantity of lexical collocations was further broken down into two sub-groups: the tokens of lexical collocations (i.e., frequency) and the types of lexical collocations (i.e., variety). ‘Token’ accounted for all the acceptable or “solid lexical collocations” (Sung, 2003, p. 49) collected from the student writing while ‘type’ included every new lexical collocation without a repetition. The concept of dividing frequency/token from variety/type was not new; however, previous studies on the
collocation production have never made a distinction between these two. The current study intended to implement this new research design, hoping to see results that are more convincing.

Last, the students’ online writing was measured and reported by the automated essay grading (i.e., AEG, Cheng, 2006) system provided by Criterion. In this study, the three sets of descriptive data (i.e., tokens of lexical collocations, types of lexical collocations, and online writing scores) were collected from the same web-based writing program. The Pearson correlation coefficient was adopted to determine the degree of a relationship between two quantifiable variables.

3.2 Study Setting and Subjects

The subjects in this study were 62 EFL students at a university of science and technology in southern Taiwan. Forty-one of them were English majors (i.e., 23 juniors and 18 seniors) and the other twenty-one were non-English majors who were grouped together regardless of their academic years or majors by a placement test as a class of the most advanced English fluency (i.e., Level Four in the subjects’ university) in the General Education. The students joined this study voluntarily because they wanted to become familiar with the latest format of English proficiency test which has made online writing an unavoidable component. They were also eager to know whether their 9–10 years of formal English training could actually prepare them for this kind of writing task. Besides, after learning the intent of the current study, the students all expressed their interest in knowing the possible relation between lexical collocations and their online writing performance.

The nature of subject selection procedure qualified this study as a quasi-experimental design. It is similar to true experimental designs, except that the participants are not randomly selected for the study (Mertens, 1997).

3.3 Instrument—Criterion (7.1 Version) Web-based Writing Program

The only data collection instrument used in this study was Criterion, ETS, a web-based writing program built upon a combination of four modern technologies: word processors, the internet, corpus, and automated essay grading systems (Cheng, 2006). This type of web-based writing programs have been widely adopted as an alternative tool in teaching EFL/ESL writing, and have shown their functions in improving students’ writing skills (See Attali, 2004; Cheng, 2006; Grimes & Warchauer, 2006; Chodorow & Burnstein, 2004 for further discussions).

Criterion was chosen based on four major concerns. First, online writing is nearly the norm for every EFL/ESL writing class. Our students almost never conduct paper-pencil writing anymore. Criterion comes closest to their regular writing setting. This study collected online written data produced naturally by our students. This type of computer-based writing is just beginning to receive attention. The current study intends to bring insights to the previous findings.

Second, the feature of e-portfolios, recording the first and the most recent essays online, allowed the researcher to access students’ writing sample easily. Immediately after the online writing test, all the student writings can be converted to text-only documents, speeding up the process of tallying lexical collocations. In addition, the e-rater function
assigned a holistic score\(^5\) (i.e., 1~6, from the lowest to the highest) to each student essay according to the pre-programmed linguistically-specific features and instant comparisons to previous corpus-based writing samples. It only took seconds to determine a student’s writing. With its 97\% accuracy rate in comparison to human scoring (Chodorow & Burnstein, 2004), Criterion gave the current and future studies of quantitative nature a chance to include and analyze a large sample of written data.

Finally, speed online writing has been included in many major ETS-designed English proficiency exams. For instance, the new TOEFL-iBT is requesting its test-takers to conduct two types of online writing—‘an independent writing’ in responding to a prompt and ‘an integrated writing’ answering to an audio or a written passage. Not for long, our EFL/ESL learners will need assistance to cope with this new type of challenge. The current study is taking a brave step ahead. By examining the correlation between online writing and frequency/variety of lexical collocations, it aims at finding out possible techniques or skills which EFL/ESL can use when taking an online writing test.

To sum up, Criterion was selected for a two-fold purpose. Its web-based features reduce the burden of grading and data collecting. It is also one of the few available web-based writing programs (e.g., currently Criterion and MyAccess) that our students can benefit from for their current needs and future academic preparation.

### 3.4 Data Collection Procedures

This study was comprised by two stages. In the first stage, the 41 English majors took the same online writing with the same topic—\textit{Dress for Success} in a computer lab within the allotted 45-minute writing time. Once the students logged in their own accounts, they saw the following prompt:

Until the 1990s, business suits were standard attire in the corporate world. These days corporate dress codes vary -- some employers still expect employees to wear suits while others welcome much more casual dress. While many people welcome the increased comfort and freedom of casual business attire, others believe that casual dress in the corporate workplace is unprofessional or even harmful. Write an essay exploring the pros and cons of the trend toward casual dress in the workplace. Explain your position with reasons and examples from your own experience, observations, or reading (Criterion, 7.1 Version, 2007).

They wrote in the word processor document provided by Criterion web-based writing program where all the writing samples were stored under each student’s account as well. The 24 non-English majors took the test one day after the English majors. They used the same computer lab and wrote their essays based on the same topic.

In the second stage, the researcher and four research assistants, from the instructor’s access account, retrieved the 65 students’ writing samples and recorded their scores awarded by the online program. Because three non-English majors’ writing scores were categorized ‘N/A,’ indicating their essays were not gradable, the researcher decided to leave out their data. Altogether, there were 62 subjects at the end of the study.

### 3.5 Tallying Lexical Collocations

---

\(^5\) For a detailed 6-point scoring rubric, please see Appendix B.
Tallying lexical collocations was done manually with caution although it was time and energy consuming. The four research assistants first underlined all possible lexical collocations recognized in the student writing according to the six subtypes\(^6\) of lexical collocations suggested by the *BBI dictionary of English word combinations* (Benson, et al., 1997) and the seventh ‘noun + noun’ subtype recognized by Hausmann (1999) and Kimmes (2004). Altogether, seven types of lexical collocations were adapted in this study.

The researcher judged whether a lexical collocation was acceptable with the help of the *BBI dictionary of English word combinations* (Benson, et al., 1997), *Oxford Collocations Dictionary for Learners of English* (Lea, et al., 2002), and three online corpora: American National Corpus (http://americannationalcorpus.org), Simple Search of British National Corpus (http://sara.natcorp.ox.ac.uk/lookup.html), and VLC Web Concordancer (http://vlc.polyu.edu.hk/concordance/WWWConcappE.htm), and two native speakers who are both experienced EFL teachers. Online corpora were utilized in this study because evidence showed that a corpus could illustrate how words collocate (Aghbar, personal communication, 2002; Schmitt, 2000; McCarthy, 2004b). The three online corpora were used because they were both free of charge. In addition, they were designed to include a wide range of samples from different registers and language domains. The three corpora were selected along with the two native speakers from the U.S. as a balance was made on the judgment of lexical collocations between British English and American English. Besides, although the purpose of the online writing was to elicit the subjects’ use of collocations, it was also used as a measurement for the subjects’ writing fluency. When an acceptable lexical collocation was found with spelling or grammatical errors, it was also counted as a valid one. It was deemed as an effort of using lexical collocations. Such mechanic deficiency would later automatically result in the deduction of the students’ writing scores determined by the web-based writing program. Nevertheless, the inclusion of misspelled lexical collocations would not harm the accuracy of tallying.

Table 3

<table>
<thead>
<tr>
<th>Subtypes and Examples of Lexical Collocations Recognized in the Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>L1 verb + noun</td>
</tr>
<tr>
<td>L2 adjective + noun</td>
</tr>
<tr>
<td>L3 noun + verb naming an action</td>
</tr>
<tr>
<td>L4 noun 1 of noun 2</td>
</tr>
<tr>
<td>L5 adverb + adjective</td>
</tr>
<tr>
<td>L6 verb + adverb</td>
</tr>
<tr>
<td>L7 noun + noun</td>
</tr>
</tbody>
</table>

\(^6\) In the *BBI Dictionary*, there are 7 subtypes of lexical collocations. L1 verb + noun (creation) and L2 verb + noun (eradication) were merged into L1 verb + noun in this study. ‘Subtype’ is also used to refer to the 7 original lexical collocations of dictionary definitions while ‘type’ in this study is reserved for describing the students’ variety in their use of lexical collocations.
At the last stage, the researcher divided the valid lexical collocations into the groups of “tokens” and “types” of lexical collocations. The data were eventually computed by Statistical Package for the Social Science (SPSS) Version 14.

4. RESULTS AND DISCUSSIONS

4.1 Analysis of Research Questions

Test results were examined to answer the two major questions for correlation (1) between the subjects’ frequency (i.e., tokens) of lexical collocations and their writing and (2) between the subjects’ variety (i.e., types) of lexical collocations and their writing. The Pearson correlation was used to describe the correlations between the 41 EFL subjects’ use of lexical collocations and their online writing scores. Table 4 lists all the correlations between every set of two variables.

<table>
<thead>
<tr>
<th></th>
<th>English Majors’ Writing Scores</th>
<th>Non-English Majors’ Writing Scores</th>
<th>All Students’ Writing Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens of Lexical Collocations</td>
<td>( r = .261 ) ( p = .100 )</td>
<td>( r = .200 ) ( p = .386 )</td>
<td>( r = .544^{**} ) ( p = .000 )</td>
</tr>
<tr>
<td>Types of Lexical Collocations</td>
<td>( r = .384^{*} ) ( p = .013 )</td>
<td>( r = .162 ) ( p = .483 )</td>
<td>( r = .809^{**} ) ( p = .000 )</td>
</tr>
<tr>
<td>Total Words</td>
<td>( r = .621^{**} ) ( p = .000 )</td>
<td>( r = .809^{**} ) ( p = .000 )</td>
<td>( r = .890^{**} ) ( p = .000 )</td>
</tr>
</tbody>
</table>

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

In terms of the correlation between all the students’ tokens of lexical collocations and their online writing scores, it is significantly and positively correlated (\( r = .544 \) at \( p < .000 \)). For the same category of correlation, no obvious relation is found if the English and non-English majors are considered separately. As for the correlation between all the students’ types of lexical collocations and their online writing scores, it is even more significantly and positively correlated (\( r = .809 \) at \( p < .000 \)), indicating that the variety of lexical collocations seemed to be a better indicator for the students’ writing scores. Also another significant correlation (\( r = .384 \) at \( p < .013 \)) can be found between the English majors’ types of lexical collocations and their online writing scores.

An unexpected but interesting significant correlation is also found between all the students’ essay length (i.e., total words) and their online writing scores. This may not be a coincidence since underdeveloped, short essays could have been graded with lower scores. This is beyond the scope of the current study. Later big scale studies should be able to address this issue.

The findings of this study is well supported by Zhang (1993) who first conducted an empirical study examining the correlation between the EFL learners’ use of lexical collocations and their writing fluency. Drawing his conclusion from paper-and-pencil writing samples, Zhang pinpointed that quantity, but more important, quality for use of
collocations distinguish between native and non-native college freshmen writing as well as between good and poor college freshmen writing. This study also found that the quality of lexical collocations in terms of variety is indicative of the quality of the online writing of Taiwanese EFL College students, both English and non-English majors.

4.2 Analysis of Seven Subtypes of Lexical Collocations

A closer look at the students’ use of seven lexical collocation subtypes from the online writing may reveal findings of values. Table 5 presents Pearson correlation for the seven lexical collocation subtypes and their relation to the students’ writing.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Pearson Correlation between Student Writings and Seven Lexical Collocation Subtypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English Majors’ Writing Scores</td>
</tr>
<tr>
<td>L1 (verb + noun)-Tokens</td>
<td>$r = -.364^*$</td>
</tr>
<tr>
<td>L1 (verb + noun)-Types</td>
<td>$r = .247$</td>
</tr>
<tr>
<td>L2 (adjective + noun)-Tokens</td>
<td>$r = -.085$</td>
</tr>
<tr>
<td>L2 (adjective + noun)-Types</td>
<td>$r = .410^{**}$</td>
</tr>
<tr>
<td>L3 (noun + verb)-Tokens</td>
<td>$r = -.149$</td>
</tr>
<tr>
<td>L3 (noun + verb)-Types</td>
<td>$r = -.149$</td>
</tr>
<tr>
<td>L4 (noun 1 of noun 2)-Tokens</td>
<td>$r = .084$</td>
</tr>
<tr>
<td>L4 (noun 1 of noun 2)-Types</td>
<td>$r = .043$</td>
</tr>
<tr>
<td>L5 (adverb + adjective)-Tokens</td>
<td>$r = .294$</td>
</tr>
<tr>
<td>L5 (adverb + adjective)-Types</td>
<td>$r = .360^*$</td>
</tr>
<tr>
<td>L6 (verb + adverb)-Tokens</td>
<td>$r = -.302$</td>
</tr>
<tr>
<td>L6 (verb + adverb)-Types</td>
<td>$r = -.292$</td>
</tr>
<tr>
<td>L7 (noun + noun)-Tokens</td>
<td>$r = .139$</td>
</tr>
<tr>
<td>L7 (noun + noun)-Types</td>
<td>$r = .026$</td>
</tr>
</tbody>
</table>

Note: ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed). (a) Cannot be computed because at least one of the variables is constant.

If examining the correlations from the perspective of the entire student group, it is evident to see positive correlations occurring between the students’ writing scores and L1 ‘verb + noun’ subtype (both tokens and types), L2 ‘adjective + noun’ subtype (both tokens and types), L4 ‘noun 1 of noun 2’ subtype (both tokens and types), and L7 ‘noun + noun’ subtype (tokens only). Since the sample in the current study is too small, no conclusions can be made based on the correlations shown in Table 5. Previous studies of English collocations have never addressed the issue of lexical collocation subtype acquisition, thus making it difficult and almost impossible to compare the findings to those of the former studies.

However, if the total types (i.e., variety) of lexical collocations used by the students in their online writing are listed according to the holistic scores they received, a pattern of lexical collocation development can be observed.
From Figure 1, it is obvious to see the L1 and L2 subtypes might have been the earliest acquired lexical collocations as they developed in two relatively paralleled and upward lines, beginning from the Score 1 to the Score 6 group. As for the other 5 subtypes, no concrete patterns can be found.

Again, the samples in this study are limited. Any claims on the developing patterns of lexical collocations can be dangerous. The findings here, however, should be considered as an initial step into later studies on the possible acquisition pattern of lexical collocations.

5. CONCLUSION

5.1 Pedagogical Implications
The study has shown that frequency and variety of lexical collocations were both significantly associated with Taiwanese university EFL learners’ online writing. And the association was even stronger between the variety of lexical collocations and the 62 student writers’ online writing scores. Future bigger scale studies might still need to be carried out in order to examine whether the use (i.e., frequency and variety) of lexical collocations will benefit EFL learners’ writing fluency, particularly assessed by web-based writing programs. Based on the current study, the first implication we can make is to suggest our student writers to learn and use a wide variety of lexical collocations as they seem to help more than merely putting as many collocations as possible.

The next pedagogical implication concerns the subtypes of lexical collocations for Taiwanese college EFL learners. The study showed that the majority of the subjects were using a very low number of L3 (noun + verb), L4 (noun 1 of noun 2), L5 (adverb + adjective), L6 (verb + adverb), and L7 (noun + noun) subtypes of lexical collocations as long as they were writing online. That is to say, Taiwanese university EFL students might not have sufficiently gained these five lexical collocation subtypes in their online written production. Based on such findings, EFL writing teachers should especially emphasize on the practice of these certain subtypes (i.e., L3~L7) in class.

5.2 Limitations of the Study

The fact that the subjects of the study were not randomly selected is perhaps one obvious limitation. The subjects were limited to the students at a national university of science and technology from southern Taiwan. Therefore, while the study revealed interesting findings about the relationship between the Taiwanese EFL university learners’ use of lexical collocations and their online writing fluency, the findings may not be easily generalized beyond the subjects of the study.

Another limitation of this study is related to the scope of subjects. With the assistance of the web-based writing program, Criterion, this study might have had a good chance to recruit more student participants than it actually had. When each student needed to pay to maintain an online writing account, it became very difficult if the study had not had the sponsorship from Chun Shin, ETS Taiwan Distributor. Future studies of the same kind will need to take budget into consideration.

5.3 Recommendations for Future Research

Studies on collocations have been carried out in the past two decades but our understanding of collocations is still not enough. As the current study explored Taiwanese college students’ use of lexical collocations as well as their online writing fluency, several recommendations can be made.

First, a strong relationship was found both between the frequency of lexical collocations and online writing fluency and between the variety of lexical collocations and online writing fluency among the Taiwanese university EFL learners in the study. However, it needs a bigger sample group to determine whether the frequency or variety of lexical collocations matters more. Thus, Taiwanese EFL learners’ use of collocations in the online writing could be further examined. In the current study, the researchers used only one type of online writing test to elicit the subjects’ use of lexical collocations. Future studies could consider using different web-based programs (e.g., MyAccess) or asking student participants to conduct a group of online essays with a wide variety of
writing prompts for eliciting the subjects’ written data. Second, the current study looked at the subjects’ use of collocations in relation to their online writing fluency. Future research could be extended to investigate the relationship between use and even knowledge of lexical collocations and other language skills to be performed on CMC tools. For instance, the TOEFL-iBT does require our students to take internet-based speaking, listening, and reading tests. These areas of computer-mediated language skills are the new territory we have not touched upon. Finally, the study also found Taiwanese EFL students used differently on the seven major categories of lexical collocations. The students were so limited in using the L3–L7 subtypes of lexical collocations in their online writing. Hence, it could be a new direction to continue investigating if there are some patterns of development in the EFL learners’ use and knowledge of lexical collocations particularly related to the seven subtypes.

Acknowledgements

I would like to thank Chun Shin Limited, ETS Taiwan Distributor, for providing this trail project to set up the two online writing courses at National Kaohsiung First University of Science and Technology in spring, 2007. I am also grateful to the supports of my students in the two writing courses and four project assistants for coordinating a wide array of tasks to make this project possible. Without their dedication, my plan to write up this paper for the whole EFL community today will not be possible.
References


Chen, P-C. (陳寶珍, 2002). A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan. Unpublished master thesis, National Taiwan Normal University, Taipei, Taiwan.


Hsu, J-Y & Hsu, L-C. (許正義&徐麗珍, 2007). Teaching lexical collocations to enhance listening comprehension of English majors in a technological university of Taiwan. Soochow Journal of Foreign Languages & Cultures, 24:1-34. Taipei, Taiwan, ROC: Soochow University.


Publications.
## Appendix A—Empirical Studies on Collocation Error Analysis

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Subjects</th>
<th>Instruments</th>
<th>Focuses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lombard (1997)</td>
<td>8 graduate students</td>
<td>1. questionnaire</td>
<td>1. lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. writing assignments</td>
<td>2. grammatical collocations</td>
</tr>
<tr>
<td>Liu (1999a)</td>
<td>128 university students</td>
<td>1. collocation test (fill-in-the-blank)</td>
<td>1. lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. grammatical collocations</td>
</tr>
<tr>
<td>Liu (1999b)</td>
<td>127 university students</td>
<td>1. final exam papers</td>
<td>1. lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. compositions</td>
<td>2. grammatical collocations</td>
</tr>
<tr>
<td>Huang (2001)</td>
<td>60 college students</td>
<td>1. collocation test (fill-in-the-blank)</td>
<td>1. free combination</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. restricted collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. figurative idioms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. pure idioms</td>
</tr>
<tr>
<td>Chen (2002)</td>
<td>30 senior high school students</td>
<td>1. three writing assignments</td>
<td>1. lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. grammatical collocations</td>
</tr>
<tr>
<td>Hsueh (2003)</td>
<td>87 senior high school students</td>
<td>1. writing assignments</td>
<td>Lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(adjective + noun)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(verb + noun)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. grammatical collocations</td>
</tr>
<tr>
<td>Wang (2005)</td>
<td>75 university students</td>
<td>1. a vocabulary knowledge test</td>
<td>5 subtypes of lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. 3 oral elicitation tasks</td>
<td>(verb + noun, noun + verb, noun 1 + noun 2, verb + adverb, adverb +</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. a questionnaire</td>
<td>adjective)</td>
</tr>
<tr>
<td>Wu (2005)</td>
<td>178 university students</td>
<td>1. collocation test (fill-in-the-blank)</td>
<td>verb + noun</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lexical collocations</td>
</tr>
<tr>
<td>Chen (2006)</td>
<td>40 junior college students</td>
<td>1. midterm exam papers</td>
<td>5 subtypes of lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. student compositions</td>
<td>(verb + noun, noun + verb, noun 1 + noun 2, verb + adverb, adverb +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>adjective)</td>
</tr>
<tr>
<td>Jeng (2006)</td>
<td>university students</td>
<td>1. student compositions</td>
<td>3 subtypes of lexical collocations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(adjective + noun, noun + verb, verb + noun)</td>
</tr>
</tbody>
</table>
Appendix B—Scoring Guide by Criterion (7.1), ETS (http://criterion.cs.ets.org)

**Score of 6:** You have put together a convincing argument. Here are some of the strengths evident in your writing:

Your essay:

- Looks at the topic from a number of angles and responds to all aspects of what you were asked to do
- Responds thoughtfully and insightfully to the issues in the topic
- Develops with a superior structure and apt reasons or examples (each one adding significantly to the reader's understanding of your view)
- Uses sentence styles and language that have impact and energy and keep the reader with you
- Demonstrates that you know the mechanics of correct sentence structure, and American English usage virtually free of errors

**Score of 5:** You have solid writing skills and something interesting to say. Look at the 6 sample essay to get ideas on how to develop your ideas more fully or use language more persuasively and consistently.

Your essay:

- Responds more effectively to some parts of the topic or task than to other parts
- Shows some depth and complexity in your thinking
- Organizes and develops your ideas with reasons and examples that are appropriate
- Uses the range of language and syntax available to you
- Uses grammar, mechanics, or sentence structure with hardly any error

**Score of 4:** Your writing is good, but you need to know how to be more persuasive and more skillful at communicating your ideas. Look at the 5 and 6 sample essays to see how you could be more persuasive and use language more effectively.

Your essay:

- Slights some parts of the task
• Treats the topic simplistically or repetitively
• Is organized adequately, but you need more fully to support your position with discussion, reasons, or examples
• Shows that you can say what you mean, but you could use language more precisely or vigorously
• Demonstrates control in terms of grammar, usage, or sentence structure, but you may have some errors

**Score of 3:** Your writing is a mix of strengths and weaknesses. Working to improve your writing will definitely earn you more satisfactory results because your writing shows promise.

In one or more of the following areas, your essay needs improvement. Your essay:

• Neglects or misinterprets important parts of the topic or task
• Lacks focus or is simplistic or confused in interpretation
• Is not organized or developed carefully from point to point
• Provides examples without explanation, or generalizations without completely supporting them
• Uses mostly simple sentences or language that does not serve your meaning
• Demonstrates errors in grammar, usage, or sentence structure

**Score of 2:** You have work to do to improve your writing skills. You probably have not addressed the topic or communicated your ideas effectively. Your writing may be difficult to understand.

In one or more of the following areas, your essay:

• Misunderstands the topic or neglects important parts of the task
• Does not coherently focus or communicate your ideas
• Is organized very weakly or doesn't develop ideas enough
• Generalizes and does not provide examples or support to make your points clear
• Uses sentences and vocabulary without control, which sometimes confuses rather than clarifies your meaning

**Score of 1:** You have much work to do in order to improve your writing skills.
You are not writing with complete understanding of the task, or you do not have much of a sense of what you need to do to write better. You need advice from a writing instructor and lots of practice.

In one or more of the following areas, your essay:

- Misunderstands the topic or doesn't show that you comprehend the task fully
- Lacks focus, logic, or coherence
- Is undeveloped--there is no elaboration of your position
- Lacks support that is relevant
- Shows poor choices in language, mechanics, usage, or sentence structure which make your writing confusing