英語為外語學習者在全民英檢寫作測驗中語塊之使用

EFL Learners' Use of Lexical Chunks on the GEPT Writing Test

by

孫晶昱 Ching-Yu Eunyce Sun

THESIS

Presented to the Faculty of the

Department of Foreign Languages and Literature of

Tunghai University

in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS in

Teaching English As A Foreign Language

TUNGHAI UNIVERSITY

June 2018

中 華 民 國 一百零七 年 六 月

東海大學碩士學位論文考試審定書

外國語文學系 碩士班

研究生 孫晶昱 所撰之論文：

EFL Learners' Use of Lexical Chunks on the GEPT Writing Test

英語為外語學習者在全民英檢寫作測驗中語塊之使用

經本委員會審查，符合碩士學位論文標準。

學位考試委員會

指導教授 [illegible]

考試委員 [illegible]

[illegible]

系主任 [illegible]

中華民國 107 年 6 月 26 日

TABLE OF CONTENT

ACKNOWLEDGEMENTS ..... **i**

ABSTRACT (English) ..... **ii**

ABSTRACT (Chinese) ..... **iv**

LIST OF TABLES ..... **vi**

LIST OF FIGURES ..... **vii**

CHAPTER 1 INTRODUCTION **1**

Statement of the Problems 3

Purpose of the Study 4

Significance of the Study 5

Definition of Terms 5

CHAPTER 2 REVIEW OF THE LITERATURE **8**

Vocabulary Knowledge in Second Language Performance 8

Classification of Lexical Chunks 12

Nattinger & DeCarrico's Classification 12

Lewis' Classification 15

Other Classifications from Corpora Research 16

Significance of Lexical Chunks in Language Use 18

Offering Processing Efficiency 19

Promoting Language Fluency 19

Enhancing Interactive Understanding 20

Improving Writing Performances 21

General English Proficiency Test in Taiwan 24

CHAPTER 3 METHOD 27

Data Collection Procedures 27

Data Analysis Procedures 28

Mini Pilot Study 34

The Results 37

The Modifications 38

CHAPTER 4 RESULTS AND DISCUSSION 41

Difference in the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners 41

Difference in the Numbers of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners 44

CHAPTER 5 CONCLUSION 51

Summary of the Major Findings 51

Difference in the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners 51

Difference in the Numbers of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners 52

Pedagogical Implications 53

Limitations of the Study 55

Suggestions for Future Research 55

REFERENCES 58

APPENDICES 63

APPENDIX A Three Writing Topics and Prompts for the Current Study 63

APPENDIX B Coding Sheet 65

APPENDIX C Coding Samples 67

APPENDIX D Intercoder Agreement 70

APPENDIX E Two Writing Topics and Prompts (Mini Pilot Study) 71

APPENDIX F Intercoder Agreement (Mini Pilot Study) 72

APPENDIX G Distribution of Five Categories of Lexical Chunks (Mini Pilot Study) 73

ACKNOWLEDGEMENTS

I would like to take this chance to express my greatest gratitude to those who helped and supported me throughout the process of writing my thesis.

First, I would like to deeply thank my thesis advisor, Dr. Kai-Lin Wu for giving me the essential support and encouragement. She not only spared no efforts guiding me to carry out the study patiently, but also gave me valuable comments. She always makes me believe that I can overcome the obstacles. Without her, the thesis could not be accomplished. I sincerely appreciate for everything she has done for me. She also sets a good example being an inspiring and supportive teacher.

Second, I would also like to thank Dr. His-Chin Chu and Dr. Mei-Hua Chen for being my committee members. With all the insightful comments they gave me, my thesis was truly improved. In addition, I would like to thank Dr. Jung-Han Chen for the instruction in statistics and the other coder, Kirsi Chen for the hard coding. I'm also grateful to LTTC (The Language Training & Testing Center) for valuable writing samples and Dr. Chia-Hui Chiu for assisting in the recommendation and application.

Last but not least, special thanks to my husband, Yung-Lung Hu, who always supports and comforts me. I am deeply grateful to him for all he has done for me and my family. Moreover, I really appreciate that my sister always diligently takes care of our parents to let me focus on the study. In the graduate school life, having my wonderful classmates' company and encouragement is the luckiest. I would like to share the joy of graduation with my dear parents, family, teachers and classmates.

Ching-Yu Eunyce Sun
Tunghai University
June 2018

EFL Learners' Use of Lexical Chunks on the GEPT Writing Test

ABSTRACT

Chunking refers to a phenomenon that human mind tends to store and retrieve the lexical phrases as whole units (Pawley & Syder, 1983). Research has confirmed the chunking phenomenon in language use by native speakers and found a significant amount of lexical chunks existing in native speakers' verbal expressions. Research has also shown that L2 speakers, compared to L1 speakers, have limited use of lexical chunks in their L2 oral production. Given that a mastery of lexical chunks is an indication of high level of language proficiency, discussions have been made on how to help L2 learners increase use of lexical chunks in order to facilitate fluency and accuracy in speaking. Few studies, however, have been carried out to investigate EFL learners' use of lexical chunks.

The proposed study aims to examine and compare use of lexical chunks by Taiwanese EFL learners who received different ratings in a writing test situation. One hundred and eighty writing samples of the GEPT intermediate level were obtained from the test developer LTTC (The Language Training & Testing Center), with 45 samples from each rating group (two to five). The researcher and another English teacher examined the samples, identified and manually classified the lexical chunks into the

following categories: (1) polywords, (2) collocations, (3) institutionalized expressions, (4) phrasal constraints, (5) sentence builders (Nattinger & DeCarrico, 1992; Michael Lewis, 1993). Comparisons were made among samples that received different ratings from two to five. One-way ANOVA was carried out to examine the difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of their use of lexical chunks.

The results showed that firstly, there is a significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of the total number and the number of different categories of lexical chunks. Secondly, by looking at each category of lexical chunks, the EFL learners in the 5-point group used more lexical chunks than other groups. Those in the 2-point group, on the other hand, used the fewest lexical chunks. Thirdly, among different ratings, phrasal constraints were the most frequently used category of lexical chunks, followed by collocations, polywords, and sentence builders. Institutionalized expressions were the least used one of the five categories. Lastly, a significant difference was found between the passing groups (4-point and 5-point) and the failing groups (2-point and 3 point). Based on the findings, pedagogical implications were discussed and future research was suggested.

Keywords: lexical chunks, GEPT writing test, EFL learners

英語為外語學習者在全民英檢寫作測驗中語塊之使用

摘要

語塊是指一種現象，即人的大腦心智傾向將詞彙短語當作一整個單位以進行存儲和檢索（Pawley＆Syder,1983）。研究證實了母語者在語言使用中的語塊現象，並且發現母語者的口語表達中存在大量的語塊。研究也顯示，第二語言學習者與母語者相較之下，在第二語言的口語表達中對語塊的使用有限。由於對語塊的掌握是語言精熟程度的一個指標，應如何幫助二語言學習者增加語塊的使用以促進口語的流利度和準確性已被注意及討論。然而，少有研究調查探討英語為外語學習者對語塊之使用。

本研究目的為調查和比較台灣英語為外語學習者在寫作考試中獲得不同級分其語塊之使用。一百八十份全民英檢中級的寫作樣本來自測驗研發機構 LTTC 語言訓練測驗中心，每個級分組別（二至五分）都有四十五份樣本。研究者和另一位英語老師將以人工方式識別和分類樣本中出現的語塊。語塊分類的依據如下：（1）多元詞，（2）搭配詞，（3）習慣用語（4）短語框架，（5）句子建構（Nattinger & DeCarrico，1992; Michael Lewis，1993）。本研究將對二到五分的樣本進行比較，並以單因子變異數分析來檢查各級分組別在使用語塊上的差異。

本研究結果顯示：首先，在 GEPT 寫作測驗中得到不同級分的學習者，在語塊

使用的總數量和不同類別的數量上，皆存在顯著差異。 其次，藉由檢視語塊的不同類別，發現 5 級分組的學習者比其他組別使用更多的語塊；而 2 級分組，不論在哪一個語塊類別，使用的數量皆為最少。 第三，在不同的級分組別中，短語框架是最常被使用的語塊類別，其次為搭配詞，多元詞和句子構建。 習慣用語是五個語塊類別中，最少被使用的。 而研究者同時也發現在寫作測驗中，及格組（4、5 級分）和不及格組（2、3 級分）之間存在顯著差異。最後，根據研究結果，提供教學和未來研究上之建議作為參考。

關鍵字: 語塊、全民英檢寫作測驗、英語為外語學習者

LIST OF TABLES

Table 2.1 Characteristics of Four Types of Lexical Chunks 14

Table 3.1 Distribution of Writing Samples 28

Table 3.2 Features and Examples of Different Types of Lexical Chunks 30

Table 4.1 Mean Frequencies and Standard Deviations of the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners 41

Table 4.2 One Way ANOVA of the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners 42

Table 4.3 Post-Hoc Tests of Mean Difference of the Total Numbers of Lexical Chunks 43

Table 4.4 Mean Frequencies and Standard Deviations of the Number of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners 45

Table 4.5 One Way ANOVA of the Numbers of Five Categories of Lexical Chunks 46

Table 4.6 Post-Hoc Tests of Mean Difference among Four Categories of Lexical Chunks 47

LIST OF FIGURES

Figure 4.1 Mean Frequencies of the Number of Five Categories

of Lexical Chunks across Four Ratings of Writing by EFL Learners……46

CHAPTER 1

INTRODUCTION

Writing is a process of using words in an organized manner to communicate thoughts and ideas in a readable form (Flower & Hayes, 1981; Graham, 2006). For many EFL learners, writing serves as the most challenging part among the four language skills—listening, speaking, reading and writing since it is much more complex and demanding than other language competence. Therefore, most of GEPT (General English Proficiency Test) test takers have been studying English for more than six years and have learned a great deal of English grammar and vocabulary; however, they still fail in the writing section. They get used to searching for English equivalents for Chinese words in their minds word by word, and then making English sentences by the grammatical rules. Besides treating grammar and vocabulary respectively in the traditional English classroom, single words also receive much emphasis in instruction while less attention is paid to lexical chunks.

Knowledge of vocabulary is clearly essential for writing, especially from the perspective of L2 writing instruction (Hyland, 2007). More studies show that vocabulary is often made up of ready-made multi-word sequence (Nattinger & DeCarrico, 1992; Cortes, 2004; Wray, 2002) Meanwhile, numerous terms have been

coined to refer to this type of sequence, but the most commonly used are lexical chunks (Schmitt, 2000). These lexical chunks are of great importance for L2 writers for at least two reasons. Firstly, advanced writers repeatedly use the lexical chunks which make their tasks easier rather than having to create each sentence word by word. Secondly, lexical chunks become the characteristic feature of fluent writing and are important for the development of writing that fits the expectation of readers in the academic field in terms of their frequent use (Li & Schmitt, 2009).

Lexical chunks are fixed or semi-fixed frequently used multi-word sequence, which are stored and retrieved automatically as whole units at the time of use (Lewis, 1993; Pawley & Syder, 1983; Wray, 1999; Schmitt, 2000). As computerized corpus-based research on lexical chunks develop rapidly, it becomes manifest that chunking is not the merely peripheral trait, but rather is prevalent and must be the essential feature of language (Shmitt, 2010). The use of lexical chunks in language production enhances the expression fluently and accurately. Consequently, the competence in using lexical chunks is an essential criterion to measure one's language level (Wray, 2002). A mastery of lexical chunks is an indication of high level of language proficiency.

Statement of the Problems

Since the 1980s, linguists' concern for lexical chunks, specifically a kind of multi-word unit, has been increasing (Cowie, 1992; Lewis, 1993; Pawley & Syder, 1983; Peters, 1983; Nattinger & DeCarrico, 1992). A large amount of lexical chunks exist in real language and they play an exceedingly crucial role in native speakers' expression; therefore, the use of lexical chunks has also been shown to be an essential measure of learners' language development (Ellis & Simpson-Vlach, 2008; Haswell, 1991; Hyland, 2008; Wray, 2002). Research has shown non-native speakers have limited use of lexical chunks in their L2 production (Granger, 1998; Foster,2001). This line of research however tended to focus on the speaking performance of college-level L2 learners.

Most of the previous studies paid more attention on the use of lexical chunks by native and non-native speakers or English majors in university, and focused on the oral proficiency or spoken language; nonetheless, little exploration has been made to investigate the use of lexical chunks by EFL learners at different proficiency levels. In spite of the importance of lexical chunks in language production and development, few research has examined the use of lexical chunks by learners of different proficiency levels. In addition, the findings for these research are mixed whether higher proficiency learners use more or fewer lexical chunks than the lower proficiency learners (Boers,

Eyckmans, Kappel, Stenger, & Demecheleer, 2006; Forsberg, 2010; Staples, Egbert, Biber, & McClair, 2013). More research is needed to investigate the use of lexical chunks by EFL learners who are at different levels of English writing proficiency.

Purpose of the Study

Little research has been carried out to see the use of lexical chunks by EFL learners who are at different levels of English writing proficiency. Thus, the proposed study attempts to help fill the research gap by investigating the lexical chunks produced by Taiwanese EFL learners who took the GEPT writing test. This study aims to examine writing samples of the GEPT, hoping to find explicit and detailed information about the test takers' use of lexical chunks, especially those who are at the passing level of writing proficiency.

This study intends to investigate if there is a correlation between Taiwanese EFL learners' use of lexical chunks and their GEPT writing ratings at intermediate-level. The results are later examined to answer the two major questions for correlation (1) between Taiwanese EFL learners' total number of lexical chunks and their writing ratings and (2) between Taiwanese EFL learners' number of different categories of lexical chunks and their writing ratings.

To reach the above-mentioned goal, the proposed study is designed to analyze the

writing samples of the GEPT test composed by Taiwanese EFL learners and seeks to answer:

1. Is there any significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of their use of lexical chunks?
2. Is there any significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of their use of different categories of lexical chunks?

Significance of the Study

It is hoped that the results of the study will provide pedagogical implications for both learning and teaching lexical chunks for writing in order to improve the writing performance on GEPT intermediate level. An investigation on the use of lexical chunks used by Taiwanese EFL learners who received different ratings on a writing test can help us gain a better understanding of the relationship between writing proficiency and use of lexical chunks.

Definition of Terms

The major terms used in the proposed study are defined as follows.

Lexical chunks

Lexical chunks refer to frequently-occurred, fixed or semi-fixed multi-words formed by meanings, which are stored and produced automatically as whole units, rather than grammatical rules. (Lewis, 1993; Nattinger & DeCarrico, 1992; Schmitt, 2000; Wray,2000) The categories of lexical chunks include polywords, collocations, institutionalized expressions, phrasal constraints, sentence builders.

Polywords

Polywords are fixed short phrases with no variability, and can be idioms (e.g., pull one's leg), phrasal verbs (e.g., run into, turn down, get along with), adverb phrases (e.g., after all, at any rate, at any time, and so on) and phrases functioning as transitions (e.g., on the other hand, to sum up). The meaning of the whole chunk may be apparent or opaque, and can or cannot be inferred from the meaning of the individual words (Nattinger & Decarico,1992; Lewis, 1993).

Collocations

Collocations refer to a string of words that co-occur in a natural text with greater than random frequency. According to Benson et al. (1997), Hausmann (1999), and Kimmes (2004), there are six types of collocations, including "verb＋noun (e.g., go shopping), adjective＋noun (e.g., heavy rain), noun + verb naming an action (e.g., bees

buzz), adverb＋adjective (e.g., sound asleep), verb + adverb (e.g., whisper softly), noun + noun (e.g., school uniform)".

Institutionalized expressions

Institutionalized expressions are sentence-length and invariable chunks (e.g., How are you? Practice makes perfect. Not yet. See you later). They are proverbs, aphorisms, formulas for social interaction that are stored as complete units in the mental lexicon. (Nattinger & Decarico,1992; Lewis, 1993).

Phrasal constraints

Phrasal constraints are short-to-medium-length phrases and allow variation of to-be-filled slots for the lexical elements (e.g., as ___ as, a___ ago, the ___er the ___er), including noun phrases (e.g., a piece of, a lot of), adjective phrases (e.g., afraid of, good at, famous for), and prepositional phrases (e.g., in the future, in history, around the world), etc.. (Nattinger & Decarico,1992).

Sentence builders

Sentence builders provide the framework for whole sentences to be constructed (e.g., I think that.../ Firstly,...Secondly, ... Finally,.../ It is suggested that.../ My point is that...). They allow variation of to-be-filled slots for the language users to express ideas (Nattinger & Decarico,1992; Lewis, 1993).

CHAPTER 2

REVIEW OF THE LITERATURE

The proposed study will examine Taiwanese EFL learners' use of lexical chunks in English writing. This chapter reviews related studies on lexical chunks. The first section reviews vocabulary knowledge in second language performance. The second section discusses classification of lexical chunks. The third section describes the significance of lexical chunks. The last section introduces General English Proficiency Test in Taiwan.

Vocabulary Knowledge in Second Language Performance

Vocabulary is an important component of language use and learning vocabulary is a crucial part of mastering a second language. Researchers have provided evidence to indicate that vocabulary knowledge contributes quite a lot to overall language success. Laufer's study (1992) revealed high correlations of .50-.75 between vocabulary size and reading. What's more, in terms of stepwise regression analysis, Laufer and Goldstein (2004) found that knowing the form-meaning link of words explained 42.6% of the total variance in participants' class grades, including language performance on reading, listening, speaking, writing, grammatical accuracy, sociolinguistic appropriateness, and language fluency.

Alderson (2005) conducted a study to explore systematically the relationships

between vocabulary knowledge and language proficiency. The checklist test and the vocabulary test battery correlate with reading at .64, listening from .61-.65, writing from .70-.79, and grammar at .64. Thus the correlation values squared indicated that vocabulary accounts for 37%-62% of the variance in the various language proficiency scores. The result showed that vocabulary has close relationships with the language skills and the correlation between vocabulary and writing is particularly high.

Vocabulary knowledge has been regarded as a basic principle of language proficiency and vocabulary is also considered as a strong predictor of overall writing performance when vocabulary scores are compared to more elaborate criterion measures of written expression like the Test of Written Language (Towl; Hammill & Larsen, 1978), the Stanford Achievement Test (SAT; Madden, Gardner, Rudman, Karlsen, & Merwin, 1978), and the Developmental Sentence Scoring System (Lee & Canter, 1971). Furthermore, Astika (1993) found, the vocabulary section obviously explained the largest amount of variance when using Jacobs, Zingraf, Wormuth, Hartiel and Hughey's (1981) ESL composition scale.

The breadth and depth of a writer's vocabulary has a direct impact on the second language performance, such as descriptiveness, accuracy, and quality of the writing. Johnson, Acevedo, and Mercado (2016) noted that accurate productive knowledge of

high-frequency word families was associated with L2 writing performance. L2 learners' lack of vocabulary causes writing difficulty and that vocabulary is one of the most important features that determine the quality of their writing. The difference between native speakers' and L2 learners' writing showed that native speakers use a wider range of vocabulary. Obviously vocabulary plays a significant role in the assessment of quality of writing.

Vocabulary is defined as knowledge of words which is considered vital for language development and acquisition. Nation (2001) puts forward vocabulary knowledge includes meaning, form, and use. In addition, he accounts for receptive vocabulary knowledge contains perceiving the form of a word and retrieving its meaning while listening or reading. On the other hand, productive vocabulary knowledge requires expressing a meaning through speaking or writing and producing the proper spoken or written word form. However, a noteworthy trait of vocabulary is that meaning and form do not always have a one-to-one correspondence. For example, die, pass away, and kick the bucket are all with the same meaning. The meanings can be presented by word combinations working as individual words (Schmitt, 2010).

Word combinations are as important as individual words. "the building blocks of language learning and communication are not grammar, function, notions, or some other

unit of planning and teaching but lexis, that is, word and word combinations" (Richards & Rodgers, 2001, 132) . Vocabulary has traditionally been regarded as individual words; however, now it has been getting obvious that much of lexis is composed of sequences of words which worked as single words, such as compounds, phrasal verbs, collocations, and idioms. It is used quite widespread for longer sequences of words to pattern together and a large number occurs in both spoken and written discourse. Some of these recur often enough to be considered as units, e.g. take good care of (Schmitt, 2000).

Considerable terms have been made to refer to this kind of word combinations such as "prefabricated patterns" (Hakuta, 1976), "lexicalized stems" (Pawley & Syder, 1983), "speech formulae" (Peters, 1983), "lexical phrases" (Nattinger & DeCarrico, 1992), "ready-made units" (Cowie, 1992), "lexical chunks" (Lewis, 1993), but lexical chunks are the most commonly used (Schmitt, 2000). Vocabulary research and instruction have been focusing on individual words not only because they have been regarded as the basic lexical units, but also since they are easier to work with than lexical chunks. Nevertheless, it is becoming increasingly clear that lexical chunks are important elements of language use (Pawley & Synder, 1983; Nattinger & DeCarrico, 1992; Wray, 2002; Schmitt & Carter, 2004).

Classification of Lexical Chunks

Classification of lexical chunks varies because linguists put forward distinct perspectives to classify lexical chunks. Among these attempts, the classifications proposed by Nattinger & DeCarrico (1992), Lewis (1993), and other classifications from corpora research are discussed in subsequent sections.

Nattinger & DeCarrico's Classification

The classifications of lexical chunks proposed by Nattinger and DeCarrico (1992) are mainly based on the fixedness of the phrases, the continuity of the sequence**,** the length and grammatical status, and the form of canonical or non-canonical. In the light of the criteria, lexical chunks are categorized into four categories: Polywords, Institutionalized expressions, Phrasal constraints, and Sentence builders.

(1) Polywords: Polywords are short phrases (prepositional phrase, verb phrases, infinitive…etc.) with no variability or lexical insertions with functions as topic shifter (e.g., *by the way, as opposed to*), summarizer (e.g., *all in all, in essence*), relator (e.g., *for that matter*), evaluator (e.g., *strictly speaking, beside the point*), fluency device (e.g., *so to speak*), disagreement marker (e.g., *hold your horses*). Polywords can be canonical or non-canonical.

(2) Institutionalized expressions: Institutionalized expressions refer to sentence-length sequence of words that have no variability, and include proverbs and formulas for social interaction. Most institutionalized expressions are canonical, invariable, and continuous. They have different pragmatic functions: greeting (e.g., *How do you do? How are you?*), narrative framer (e.g., *Once upon a time…and they lived happily ever after*), parting (e.g., *Have a nice day! See you later.*), advice (e.g., *A watched pot never boils.*).

(3) Phrasal constraints: Phrasal constraints refer to short-to-medium-length phrases that allow variations of phrasal and lexical categories, including NP (noun phrase), VP (verb phrase), Adj. P (adjective phrase), Adv. P (adverb phrase), and so forth (e.g., *a day/year/long time ago, as busy as a bee, the more the better, the lazier the poorer*). The phrasal constrains can be canonical or non-canonical. Most of them are continuous and have various functions.

(4) Sentence builders: The relatively long chunks have one or several slots for parameters or arguments to fill in and allow considerable variation of phrasal (NP, VP) and clausal (S) elements for expression of entire idea. They can be canonical or non-canonical, and they are both continuous and discontinuous. Comparison with other three categories, sentence builders include more discontinuous lexical phrases. For

example, not only…, but also… (e.g., *Not only is he an astronaut, but also his cousin is an astronaut.*), the Adj+er …, the Adj+er … (e.g., *The less fast food you eat, the healthier you will be.*), and I think (that) … (e.g., *I think that it's a good idea.*).

Nattinger and DeCarrico (1992) also declare that polywords are similar to institutionalized expressions though the latter belongs to the sentence level. Both of them share the same characteristics which are relatively simple, fixed as well as continuous. Nevertheless, phrasal constraints and sentence builders are much more complicated and variable as a result of both of them involve slots.

In summary, the characteristics of the four types of lexical chunks are shown in the following table (Nattinger & DeCarrico, 1992, 45).

Table 2.1

Characteristics of Four Types of Lexical Chunks

	Grammatical level	Canonical/ Non-canonical	Fixed/ Variable	Continuous/ Discontinuous
Polywords	word level	both	fixed	continuous
Institutionalized expressions	sentence level	canonical	fixed	continuous
Phrasal constraints	word level	both	somewhat variable	mostly continuous
Sentence builders	sentence level	canonical	highly variable	often discontinuous

Lewis' Classification

Lewis's classification of lexical chunks (1993) gives examples of types of lexical chunks that are beneficial for English as second language learners and clearly useful for pedagogical purposes. He proposed four types of lexical chunks: polywords, collocations, institutionalized utterances, and sentence frames or heads.

(1) Polywords: Polywords are comprised of two or three words, including transition phrases and phrasal verbs (two-and three-part verbs, like verb+ preposition). Some of the chunks show the meaning literally (e.g., *by the way, the day after tomorrow*), but some of them are totally different from the component words (e.g., *put off, all at once, look up*).

(2) Collocations: Collocations are words that co-occur with others. They have different form combinations, for instance, adjective + noun (e.g., *a short-term strategy*), verb + noun (e.g., *take medicine)*, fixed sequence (e.g., *rancid butter*), and so forth.

(3) Institutionalized utterances: Institutionalized utterances are conventionalized whole units that mainly occur in spoken language. The chunks may be full sentences with no variation but always with instantly identifiable pragmatic meaning (e.g., *Not yet./Just a moment, please./ It's nothing to do with me./I'll give you a ring./ If I were you, I'd wait.)*

(4) Sentence heads or frames: This type of lexical chunks works as the framework builder of the whole sentence. They are almost the same as institutionalized utterances; however, sentence heads or frames are often used in the written mode to structure long passages of text (e.g., *It is suggested that…, The fact is…, My point is that…, Firstly…, secondly…, finally…,* etc.) This type helps EFL learners to write fluently and accurately.

Nattinger and DeCarrico (1992) argue collocations are strings of specific lexical items without particular pragmatic functions, which co-occur with mutual expectancy greater than chance. Therefore, they assert that lexical chunks are collocations with pragmatic functions; however, collocations are apparently contained in Lewis' (1993) classification of lexical chunks. According to his viewpoint, collocations are similar to individual words but distinct from institutionalized utterances. They are usually connected with the content of what the language users express, rather than what the language users are doing, such as complaining, explaining, apologizing, etc. Since lexical chunks are fixed or semi-fixed frequently used phrases, which are stored and retrieved automatically as whole units at the time of use, collocations should also be covered in lexical chunks.

Other Classifications from Corpora Research

Biber et al. (1999) identified lexical chunks as "word combinations that recur most

commonly in a given register" from corpora research perspective (P. 184). The chunks can be considered as potential lexical chunks are merely continuous lexical sequences. Their model in the Longman Grammar of Spoken and Written English which contained the structure of NP-based, PP-based, and VP-based lexical chunks.

Cortes (2004) investigated lexical chunks in published and student writing on biology and history. The structural comparison between target lexical chunks in published biology and history writing manifests that in history there are merely two structural types which contains noun phrases and prepositional phrases. Nevertheless, in biology, there are a variety of types which comprehend noun phrases, prepositional phrases, it + be verb + adjective clause fragments, be verb + complement clause fragments and noun phrases + verb + complement clause fragments.

By comparison, Hyland's (2008) study is more aspiring. He inspected clusters in three corpora of written texts: published research articles, PhD dissertation and MA/MSc theses in four disciplines. They involved a wide cross-section academic practice from electrical engineering (EE), business studies (BS), applied linguistics (AL) to microbiology (Bio). The results revealed the most frequent patterns in the corpus were noun and prepositional phrases. Moreover, prepositional phrases were superior in PhD dissertation. Research articles also comprised significant of-phrase where they

post-modified noun phrases overwhelmingly. The master's theses covered a large number of noun phrases with prepositional phrases; however, they also significantly use the patterns to express logical or locative relations, present graphical information or emphasize the observation of a study.

A conclusion can be drawn from the aforementioned classifications. There are still no fixed criteria for classification of lexical chunks, and researchers set their own standard for their own purpose of the research. Nevertheless, when researchers use these criteria, they have to be conscious that all of these lexical chunks range between two extremes from absolutely fixed to highly free. Consequently, there are no boundaries among these types, and sometimes it is not easy to identify them clearly.

Significance of Lexical Chunks in Language Use

Shmitt (2010) pointed out that there is a good psycholinguistic basis for believing that the mind stores and processes these chunks as individual whole units and lexical chunks reduce cognitive burdens of language users when producing and comprehending language. Furthermore, lexical chunks facilitate language use because they offer processing efficiency, promote language fluency, enhance interactive understanding and improve writing performances. The significance of lexical chunks in language use is elaborated as follows.

Offering Processing Efficiency

The consensus among those who have studied lexical chunks in language use seems to be that their primary value is in alleviating the burden of mental processing. Lexical chunks are stored and retrieved as wholes, which facilitates faster processing. This means lexical chunks can provide prefabricated framework to express ideas so that speakers do not have to go through the labor of language production word by word each time they need to say or write something. Furthermore, the function which increases language production speed and fluency is called processing short-cut by Wray and Perkins (2000). They proposed that learners reduce burdens on the memory by keeping information inside lexical chunks so that it is easy to be retrieved from the memory as a whole, which facilitate saving effort in processing. Therefore, native speakers tend to call on a vast repertoire of ready-made language in their mental lexicons rather than scratch each time they speak or write. In other words, lexical chunks are memorized as units to be processed faster. They can be retrieved more easily than sequences of words which are generated creatively because of the efficient processing (Kuiper & Haggo, 1984; Pawley & Syder, 1983).

Promoting Language Fluency

Storing and retrieving ready-made lexical chunks facilitates greater fluency in

language production. If a speaker can pull the chunks readily from memory as wholes, fluency is enhanced. This lessens the amount of planning, processing, and encoding needed and gives the speaker time to concentrate on other tasks necessary, like producing specific lexical items, and planning the next unit of discourse (Wood, 2002). According to Pawley & Syder (1983), native-like fluency means native speakers have the ability to produce long strings of speech which surpass their capacity for encoding and decoding speech. It connects with language production and is the ability to link units of language with facility. In addition, Lewis (1997) proposed that fluency depends on the acquisition of a large store of fixed and semi-fixed preformed items. He suggests that lexical chunks provide the readily framework for language production, thus enhance the fluency.

Enhancing Interactive Understanding

Wray and Perkins (2000) demonstrated the socio-interactional functions for lexical chunks, which associate with facets of how people want others to treat or view them. Bygate (1988) found a wide range of syntactic and pragmatic uses of lexical chunks used in a noteworthy range of conversational functional contexts and for a wide variety of pragmatic purposes. Nattinger and DeCarrico (1992) argued that lexical chunks are the major markers of social interactions, topics, discourses and fluency devices.

Accordingly, lexical chunks highlight the direction of language production, which certainly makes the discourse more comprehensible.

Improving Writing Performances

Lexical chunks play a significant role in non-native speakers' writing proficiency (Cowie & Howarth, 1996). Haswell (1991) noted that L2 learners are required to master the use of lexical chunks in order to be successful academic writers. Nattinger and Decarrico (1992) examined the ways that lexical chunks are organized in written discourse. They found that the input of these lexical chunks can help EFL learners to express themselves well in the writing. Another study by Snellings, Van Gelderen, and de Glopper (2004) found the effects of lexical chunks on improving narrative L2 writing.

To investigate the effect of using lexical chunks on the achievement of third-year-university students of English in the descriptive essay writing, Qader (2016) carried out an empirical study. The experimental group received six sessions of treatment. The result indicated that raising students' awareness of lexical chunks was more effective than the commonly used method which focuses on form rather than meaning. This means that lexical chunks play an essential role in improving the college students' English essay writing. In conclusion, the aforementioned studies show that the

importance of the lexical chunks in developing writing skills and improving writing performance.

Many attempts have been made by researchers to explore on the relationship between the use of lexical chunks and EFL learners' writing proficiency (Cowie & Howarth, 1996; Granger 1998; Haswell, 1991; Nattinger & Decarrico, 1992; Lewis, 1997; Snellings, Van Gelderen, and de Glopper, 2004; Qader, 2016). The major findings show arousing the awareness of lexical chunks and mastering the use of lexical chunks facilitate EFL learners' writing performance. Few studies (Ferris, 1994; Hsu, 2007; Zhu, 2013; Zonghui, 2016), however, have been conducted to examine and compare lexical chunk use with different EFL learner levels, comprising the quantity and types.

Hsu (2007) analyzed the correlation between the use of English lexical collocations and the online writing of Taiwanese college English majors and non-English majors. Writing from 41 English majors and 21 non-English majors was analyzed. The results indicated a significant correlation between the Taiwanese college EFL learners' frequency (tokens) of lexical collocation and their online writing scores. In addition, a significant correlation between the subjects' variety (types) of lexical collocations and their online writing scores. The findings suggested that the variety of lexical

collocations in comparison with the frequency of lexical collocations was a better indicator for the students' writing scores.

Ferris (1994) analyzed 28 lexical variables (e.g. prepositional phrases) and syntactic variables (e.g. relative clauses) in 160 ESL compositions written by students of different language proficiencies. The results showed that the ESL students at higher levels of L2 proficiency used more of the lexical choices and syntactic constructions in their compositions than did those at lower levels. In a similar vein, Zhu's study (2013) on prefabricated chunks used by second language learners of different levels also found that L2 learners' ability in lexical chunk use is correlated with their language level. The learners with high English scores used more and better lexical chunks than learners with the low English scores.

Similarly, Zonghui (2016) conducted a study to investigate the use of lexical chunks in the English writing by Chinese college students. The study found that the juniors who passed TEM4 (Test for English Majors Band 4, which is the most widely accepted and authoritative test for English majors in China) have higher frequency of using lexical chunks than the sophomores who were preparing to pass TEM 4 in four months.

Biber et al. (2013) conducted a study on the use of lexical bundles in written

responses across three proficiency levels in the TOEFL iBT. The corpus used for this study comprised two written texts from 480 participants for a total of 960 texts and 249,417 words. For the aims of the study, they defined participant proficiency as the mean of the ETS scores on each participant's two written tasks. The corpus is further subdivided into three proficiency levels (low, medium, and high). Four-word lexical bundles were classified into prompt bundles and non-prompt lexical bundles and analyzed separately. The result indicated that lower level learners used more bundles overall but also more bundles identical to those in the prompts.

General English Proficiency Test in Taiwan

The General English Proficiency Test (GEPT) is developed by the Language Training and Testing Center (LTTC) commissioned by the Ministry of Education in Taiwan. GEPT has been administered since 2002. It is currently the most credible English proficiency test in Taiwan and often regarded as a proof of English proficiency while people are looking for a job or applying for a college. GEPT writing tests come in four levels, comprising basic, intermediate, high-intermediate and advanced. Writing tasks of different levels have different requirement on the length of the writing. For example, the intermediate-level writing, which is considered a level that Taiwanese senior high school students are expected to pass. It requires test takers to produce an

expository composition of about 120 words in one or more paragraphs in response to the writing prompt.

The composition is assessed holistically based on content, organization, grammar, word choice, and punctuation. The scores range from 0 to 5, with 5 indicating "full competence in writing", 4 "fair competence in writing", 3 "limited competence in writing", 2 "little competence in writing", 1 "lacks competence in writing", and 0 "No answer or Non-ratable". Passing grades are four or above. (https://www.lttc.ntu.edu.tw/E_LTTC/E_GEPT/intermediate.htm)

It should be noticed that it is arduous to accurately compare the before-mentioned studies since different definitions and classifications of lexical chunks were used. Notwithstanding the importance of lexical chunks in language use, few studies have explored the use of lexical chunks by learners of different proficiency levels. What's more, the findings for these studies are mixed, particularly the variety (types) of lexical chunks (e.g. Hsu, J. Y., 2007; Biber et al., 2013). Additionally, the previous studies focused more on college students' writing. More research needs to be conducted on the correlations between the use of L2 learners' lexical chunks and writing performances. Therefore, this study aims to examine and compare the use of lexical chunks by Taiwanese EFL learners who received different ratings in the writing section of GEPT

intermediate level.

CHAPTER 3

METHOD

This chapter describes the methodology of the study. It includes data collection procedures, data analysis procedures, and mini pilot study.

Data Collection Procedures

One hundred and eighty writing samples of the GEPT intermediate level were obtained from the test developer LTTC (The Language Training & Testing Center), with 45 samples from each rating group (two to five). These samples were written on the following three topics:

Topic 1: Many people have idols. In addition to film and television celebrities, athletes or writers, etc., there may be the people around. Please write an article to (1) describe the idols you like now or before; (2) explain the reasons you like.

Topic 2: In history, there are many inventions that have changed the life of mankind. Please write an article to (1) describe you want the most to see an invention in the future; (2) explain the importance of this new invention.

Topic 3: Music is an indispensable part of the lives for many people. Please write an article to (1) explain the importance of music to you and state the reasons; (2) describe your own favorite music type.

The aforementioned three writing prompts were written in Chinese (see Appendix A). Sixty samples written in response to each of the three topics will be examined. Among the sixty samples, 15 were rated five points; 15, four points; 15, three points; 15, two points. A total of 180 writing samples of different ratings were collected and then analyzed. Table 3.1 shows the distribution of the three topics and four different ratings.

Table 3.1
Distribution of Writing Samples

	5-point	4-point	3-point	2-point	Quantity
Topic 1	15	15	15	15	60
Topic 2	15	15	15	15	60
Topic 3	15	15	15	15	60
Sum	45	45	45	45	180

As shown in the table, writing samples rated 1 point are not included because the sample size, according to the LTTC, is too small for valid comparison.

Data Analysis Procedures

To analyze the 180 writing samples across different ratings, the researcher and another English teacher adopted the classification system based on Nattinger & DeCarrico and Michael Lewis to manually code and tally each lexical chunk from each sample. To include the lexical chunks proposed by Nattinger & DeCarrico and Michael

Lewis, the researcher combined their classification frameworks into a five-category framework: polywords, collocations, institutionalized expressions, phrasal constraints, and sentence builders.

(1) **Polywords** are fixed short phrases with no variability, and can be idioms, phrasal verbs, and phrases functioning as transitions (e.g., after all, all at once, put off).

(2) **Collocations** refer to a string of words that co-occur in a natural text with greater than random frequency. According to Benson et al. (1997), Hausmann (1999), and Kimmes (2004), there are six types of collocations, including "verb＋noun (e.g., go shopping), adjective＋noun (e.g., heavy rain), noun + verb naming an action (e.g., bees buzz), adverb＋adjective (e.g., sound asleep), verb + adverb (e.g., whisper softly), noun + noun (e.g., school uniform)".

(3) **Institutionalized expressions** are sentence-length and invariable chunks (e.g., How are you? Practice makes perfect.). They are proverbs, aphorisms, formulas for social interaction that are stored as complete units in the mental lexicon. While polywords and institutionalized expressions are both fixed lexical chunks with little room for variation; the former appear at the word level, and the latter at the sentence level.

(4) **Phrasal constraints** are short-to-medium-length phrases (e.g., a ____ ago,

____ as well as _____). They allow variation of to-be-filled slots for the lexical elements.

(5) **Sentence builders** provide the framework for whole sentences to be constructed (e.g., I think that…/ Firstly,…Secondly, … Finally,…). They allow variation of to-be-filled slots for the language users to express ideas. Although both phrasal constraints and sentence builders contain slots to be filled with other lexical items, the former appear at the word level, and the latter at the sentence level.

Table 3.2

Features and Examples of Different Types of Lexical Chunks

Classifications	length	variation	Examples
1. Polywords	phrase level	fixed	after all, by the way, on the other hand, put off, all at once, …
2. Collocations	phrase level	fixed or variable (productive)	
❶ verb＋noun			go shopping, play basketball,…
❷ adjective＋noun			strong tea, dirty words, heavy rain,…
❸ noun + verb naming an action			bomb explodes, bees buzz,…
❹ adverb＋adjective			sound asleep, fully aware,…
❺ verb + adverb			whisper softly, argue heatedly, …
❻ noun + noun			dress code, entrance exam,…
3. Institutionalized expressions	sentence level	fixed	How are you? Seeing is believing. Birds of a feather flock together.
4. Phrasal constraints	phrase level	variable (with slots)	a ____ ago, the ___er the ____er, _____as well as____, both ____and _____, a_____ of ____
5. Sentence builders	sentence level	variable (with slots)	I think that…/ not only…, but also…/ firstly,…secondly,…finally,…

First, a training session including the coding sheet (see Appendix B) and the coding samples (see Appendix C) was provided for the researcher and another English teacher to fully understand the classification system. In order to verify the classification done by the researcher (coder A), another English teacher (coder B) who has passed GEPT high-intermediate-level test was invited to do the classification of the lexical chunks in the writing samples. Next, coding was conducted by the two coders separately. Lastly, a discussion session was carried out to resolve any discrepancies in the coding of the two coders.

Moreover, the two coders followed the guiding principles for coding. Firstly, the lexical chunks that are misspelled or misused, are not coded as lexical chunks. Secondly, the same lexical chunks that appear multiple times were counted for only one time. For example, the lexical chunk "*In my opinion,…*" may be used three times in one composition. In such a case, they are regarded as one lexical chunk instead of three. Last, if the sentence comprises not only one type of lexical chunks, the coders would classify them into the different categorizations. For example, *She works as hard as he. Work hard* is coded as collocation (verb + adverb), and *as___as* is coded as phrasal constraints.

Whenever the two coders feel uncertain about the identification, they consult two

online corpora: Corpus of Contemporary American English (COCA) and Just the Word (JTW). The two online corpora are chosen for the study because American English and British English are the currently mainstreams for English learning in Taiwan, and then to use COCA and JTW can make a balanced judgment on lexical chunks between British and American English.

COCA（http://corpus.byu.edu/coca/） is the largest online corpus of English which is open to use without any charge. The corpus contains more than 520 million words in 220,225 texts, including 20 million words each year from 1990-2015, and it is equally sampled from transcribed spoken discourse, fiction, popular magazines, newspapers, and academic texts. It represents the American variety of English and its size is vastly larger than any other available American English corpus.

JTW (http://www.just-the-word.com/) is a very quick and easy website that directly gives collocations for a search word without the concordance lines. It shows results by part of speech (POS) and graph bars give an indication of the t-score strength. Results are based on an 80 million word subset of the British National Corpus (BNC).

For example, “dancing diva” appeared for several times in the writing samples, and the two coders felt uncertain about the combination “dancing diva” so they consulted

the two online corpora. Since COCA displayed the frequency: 1/237 and JTW showed the frequency: none, “dancing diva” was not coded as a lexical chunk in the current study.

The researcher and another English teacher served coders for the present study (coder A and coder B). They followed the classification procedure and used the pre-established classification framework to analyze 180 samples. Then, the researcher calculated the intercoder agreement rate that indicated the degree to which two coders agree on assigning categories (Brown, 2001). The intercoder agreement rate of the present study is 89% (see Appendix D).

The coding data were eventually computed using Statistical Package for the Social Science (SPSS). For research question one, One-way ANOVA (alpha decision level is at $\alpha < 0.1$) and Post-hoc analysis using a Tukey HSD were performed to examine the total numbers of lexical chunks used by Taiwanese EFL learners who received different ratings in a writing test situation of the GEPT intermediate level. For research question two, One-way ANOVA (alpha decision level is at $\alpha < 0.1$) and Post-hoc analysis using a Tukey HSD were employed to compare the numbers of different-category lexical chunks used by Taiwanese EFL learners who received different ratings in a writing test situation of the GEPT intermediate level.

Mini Pilot Study

A mini pilot study was carried out to see if the classification system of lexical chunks is feasible. Seven writing samples were retrieved from the GEPT intermediate level writing corpus provided by LTTC (https://www.lttc.ntu.edu.tw/geptscoreremark/icomposition.pdf). Among the seven samples, two were rated five points; two, four points; two, three points; and one, two points. These samples were written in response to the following two topics:

Topic 1: Generally speaking, parents often reward their children when they are doing well. Please write an article explaining (1) What do your parents usually reward you with when you are doing well? Do you think these methods are effective and appropriate? (2) Someday you are a parent, will you use the same reward method?

Topic 2: Your friend Ted, who lives in the U.S.A. recently transferred to school, is worrying about making friends at the new school. Please write an e-mail to him. The content of the letter must contain the following two points: (1) Share your experience in adapting to the new environment. (2) Suggest some ways to make new friends.

The aforementioned two writing prompts were written in Chinese (see Appendix E). In the first step of the coding procedure, the researcher underlined or circled all possible lexical chunks recognized in the writing samples based on the pre-established

classification framework proposed by Nattinger and DeCarrico (1992) and Lewis (1993). According to the framework, lexical chunks are classified into five types: polywords, collocations, institutionalized expression, phrasal constraints, and sentence builders.

(1) **Polywords** are fixed short phrases with no variability, and can be idioms, phrasal verbs, and phrases functioning as transitions (e.g., after all, all at once, put off).

(2) **Collocations** refer to a string of words that co-occur in a natural text with greater than random frequency. According to Benson et al. (1997), Hausmann (1999), and Kimmes (2004), there are six types of collocations, including "verb＋noun (e.g., go shopping), adjective＋noun (e.g., heavy rain), noun + verb naming an action (e.g., bees buzz), adverb＋adjective (e.g., sound asleep), verb + adverb (e.g., whisper softly), noun + noun (e.g., school uniform)".

(3) **Institutionalized expressions** are sentence-length and invariable chunks (e.g., How are you? Practice makes perfect.). They are proverbs, aphorisms, formulas for social interaction that are stored as complete units in the mental lexicon. While polywords and institutionalized expressions are both fixed lexical chunks with little room for variation; the former appear at the word level, and the latter at the sentence level.

(4) **Phrasal constraints** are short-to-medium-length phrases (e.g., a ____ ago,

____ as well as _____). They allow variation of to-be-filled slots for the lexical elements.

(5) **Sentence builders** provide the framework for whole sentences to be constructed (e.g., I think that…/ Firstly,…Secondly, … Finally,…). They allow variation of to-be-filled slots for the language users to express ideas. Although both phrasal constraints and sentence builders contain slots to be filled with other lexical items, the former appear at the word level, and the latter at the sentence level.

Whenever the researcher felt uncertain about the categorization, she consulted two online corpora: British National Corpus (BNC) (http://www.natcorp.ox.ac.uk/) and Just the Word (JTW) (http://www.just-the-word.com/). In the next step, in order to verify the classification done by the researcher, a second coder, who is an English teacher and has passed high-intermediate level of GEPT, was invited to do the classification of the lexical chunks appearing in the seven writing samples. The researcher provided the training in coding, including the classification of lexical chunks, using of the two online corpora and the method for coding. The second coder then did the classification independently.

In the last step, for the lexical chunks that were categorized into different types or for any lexical chunk that was considered a chunk by one coder but not by the other, a

discussion session was conducted to resolve the discrepancies in their decisions.

The Results

The researcher and another English teacher served coders for the present study (coder A and coder B). They followed the classification procedure and used the pre-established classification framework to analyze the seven samples. Then, the researcher calculated the intercoder agreement rate that indicated the degree to which two coders agree on assigning categories (Brown, 2001). The intercoder agreement rate of the mini pilot study is 92% (see Appendix F).

The results of the pilot study showed that the total number of lexical chunks used in the 4-point and 5-point (passing grades) writing samples are 2.63 times (19.25 vs. 7.33) more than the 3-point and 2-point (failing grades) writing samples (see Appendix G). The total number of lexical chunks used in the writing samples seemed to be a good indicator to distinguish writing performances that reached the passing score and those that did not on the intermediate-level GEPT.

With regard to the number of lexical chunks in different categories, the preliminary analysis shows that the writing samples with passing grades differed from the writing samples with failing grades, particularly in the categories of collocations, phrasal constraints, and sentence builders. The number of collocations used in the 4-point &

5-point writing samples are 4.6 times (10.75 vs. 2.33) more than the 3-point & 2-point writing samples. The number of phrasal constraints used in the 4-point & 5-point writing samples are 9 times (3.00 vs. 0.33) more than the 3-point & 2-point writing samples. The number of sentence builders used in the 4-point & 5-point writing samples are 1.5 times (3.00 vs. 2.00) the 3-point & 2-point writing samples.

As for other categories of lexical chunks, there is no clear difference among the seven writing samples. The number of institutionalized expressions in the writing samples with passing grades and failing grades are 0.25 & 0.33. The institutionalized expressions that the test takers used are *Hard work brings success.* (the 4-point sample) and *How are you?* (the 3-point sample). Additionally, the number of polywords in the writing samples with passing grades is close to the writing samples with failing grades (2.25 vs. 2.33).

The Modifications

After conducting the mini pilot study, the researcher clarified the classification of lexical chunks as follows. Firstly, during the coding procedure of the mini pilot study, the two coders found **one more feature** of polywords. In addition to idioms, phrasal verbs, and transitional words, polywords includes the phrases that act as an adverb (adverb phrases), for example, **after all, at any rate, at any time, and so on**.

Secondly, some lexical chunks occur as high-frequency phrases, including noun phrase (e.g., a piece of, a lot of), adjective phrases (e.g., afraid of, good at, famous for), and prepositional phrases (e.g., in the future, in history, around the world). They should be classified into "phrasal constraints" according to Nattinger and DeCarrico (1992). Therefore, the modifications of the classifications of lexical chunks are seen as follows.

Polywords are fixed short phrases with no variability, and can be idioms (e.g., pull one's leg), phrasal verbs (e.g., run into, turn down, get along with), **adverb phrases (e.g., after all, at any rate, at any time, and so on)** and phrases functioning as transitions (e.g., on the other hand, to sum up). The meaning of the whole chunk may be apparent or opaque, and can or cannot be inferred from the meaning of the individual words.

Phrasal constraints are short-to-medium-length phrases and allow variation of to-be-filled slots for the lexical elements (e.g., as ___ as, a___ ago, the ___er the ___er), **including noun phrases (e.g., a piece of, a lot of), adjective phrases (e.g., afraid of, good at, famous for), and prepositional phrases (e.g., in the future, in history, around the world), etc.**

The results from the mini pilot study seem to indicate that there were apparent differences between the passing grades (4 & 5 points) and failing grades (2 & 3 points)

in the total number and the number of collocations, phrasal constraints and sentence builders. Nevertheless, the number of polywords and institutionalized expressions used in the passing writing samples are close to those used in the failing writing samples. As a result of very small sample size for the mini pilot study, the researcher will probe into the result and find out the correlation between the use of lexical chunks and the different ratings in the writing section of the GEPT Intermediate Level by increasing the sample size.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents and discusses the results of the study. It includes the following two sections: (1) difference in the total numbers of lexical chunks used by Taiwanese EFL learners who received different ratings on the GEPT writing test; (2) difference in the numbers of different categories of lexical chunks used by Taiwanese EFL learners who received different ratings.

Difference in the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners

The total numbers of lexical chunks used by EFL learners who received in different ratings was calculated to investigate whether any significant difference exists among EFL learners who received different ratings on the GEPT intermediate-level writing test. The mean frequencies and standard deviations of the total numbers of lexical chunks used by the 180 learners on their writing tests are shown in Table 4.1

Table 4.1

Mean Frequencies and Standard Deviations of the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners

Ratings	N	Mean	Std. Deviation
2	45	3.20	1.93
3	45	5.18	2.68
4	45	10.84	3.63
5	45	14.07	4.43
Total	180	8.32	5.45

As show in the table, the EFL learners with better writing performances on the GEPT test used more lexical chunks in their writing. The mean frequencies of lexical chunks used by learners who received ratings are 3.20 (2-point group), 5.18 (3-point group), 10.84 (4-point group), and 14.07 (5-point group). To investigate if any significant difference exists among the four groups of EFL learners, One-way ANOVA was performed on the data.

Table 4.2 presents the result of One-way ANOVA of the total number of lexical chunks used by Taiwanese EFL learners who received different ratings on the GEPT writing test, including degree of freedom (df), F (MS_{BG}/MS_{WG}).

Table 4.2

One Way ANOVA of the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners

	Sum of Squares	df	Mean Square	F
Between Groups	3396.822	3	1132.274	103.550*
Within Groups	1924.489	176	10.935	
Total	5321.311	179		

Note: * Significant at $p< .01$

As shown in the table, there is a significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of the total numbers of lexical chunks (F=103.550, $p<.01$). Accordingly, a post-hoc analysis using a

Tukey HSD test was performed on the data to examine whether any significant between-group difference exists by comparing all paired means. Table 4.3 presents the result of post-hoc tests of mean difference of the total numbers of lexical chunks among different ratings, including mean difference and significance (Sig.).

Table 4.3

Post-Hoc Tests of Mean Difference of the Total Numbers of Lexical Chunks

(I) Ratings	(J) Ratings	Mean Difference (I-J)	Sig.
2	3	-1.978	.026
	4	-7.644*	.000
	5	-10.867*	.000
3	4	-5.667*	.000
	5	-8.889*	.000
4	5	-3.222*	.000

Note: * Significant at $p < .01$

As shown in the table, a significant between-group difference exists in the mean frequencies of the total numbers of lexical chunks ($p<.01$) except the comparison between 2-point group and 3-point group. This is probably because the test takers in these two failing groups used fewer lexical chunks and their means of the total numbers of lexical chunks are not significantly different from each other.

In line with previous studies (Ferris, 1994; Hsu, 2007; Zhu, 2013; Zonghui, 2016),

the present study found that the better the learners performed on the writing tests, the more lexical chunks they used in their writing. However, Biber et al. (2013) had slightly different finings. Their study on the use of lexical bundles in written responses across three proficiency levels in the TOEFL iBT found that the lowest level participants used more bundles than higher level learners. They found these participants used a higher percentage of bundles that were directly taken from the writing prompts and passages. Different from the TOEFL prompts, the GEPT writing prompts are written in Chinese so the test takers have to write the composition in English based on the writing instruction; however, the writing prompts of TOEFL iBT are in English and the lower level learners are inclined to rely on the prompt-based bundles in English. In addition, the prompt for the first task in the writing section of the TOEFL iBT is a passage about 300 words long, which certainly provides the lowest level participants with the direct source of lexical bundles. For the current study, since the instruction and prompts of the GEPT are showed in Chinese, there is no such problem.

Difference in the Numbers of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners

To further compare lexical chunks used among the different ratings, the mean and standard deviations of frequencies of five categories of lexical chunks used by learners

in the GEPT writing test were displayed as Table 4.4 and Figure 4.1.

Table 4.4

Mean Frequencies and Standard Deviations of the Number of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners

Categories / Ratings	Po	Co	Ie	Pc	Sb
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
2	0.78 (0.88)	0.93 (0.92)	0.02 (0.15)	0.96 (0.90)	0.51(0.76)
3	1.24 (1.15)	1.53 (1.47)	0.02 (0.15)	1.64 (1.35)	0.73 (0.65)
4	2.20 (1.67)	3.40 (2.05)	0.16 (0.52)	3.91 (2.75)	1.18 (0.94)
5	2.93 (1.70)	4.47 (2.59)	0.22 (0.42)	4.91 (2.81)	1.53 (1.25)
Total	1.79 (1.62)	2.58 (2.33)	0.11 (0.36)	2.86 (2.66)	0.99 (1.00)

Note: Po=Polywords Co=Collocations Ie=Institutionalized expressions Pc=Phrasal constraints Sb=Sentence builders

As shown in Table 4.4, phrasal constraints (M=2.86) were the most frequently used category of lexical chunks, followed by collocations (M=2.58), polywords (M= 1.79), and sentence builders (M=0.99). Institutionalized expressions (M=0.11) were least used one among five categories. Obviously, the EFL learners tended to use more phrasal constraints and collocations in their writing. Figure 4.1 shows this tendency.

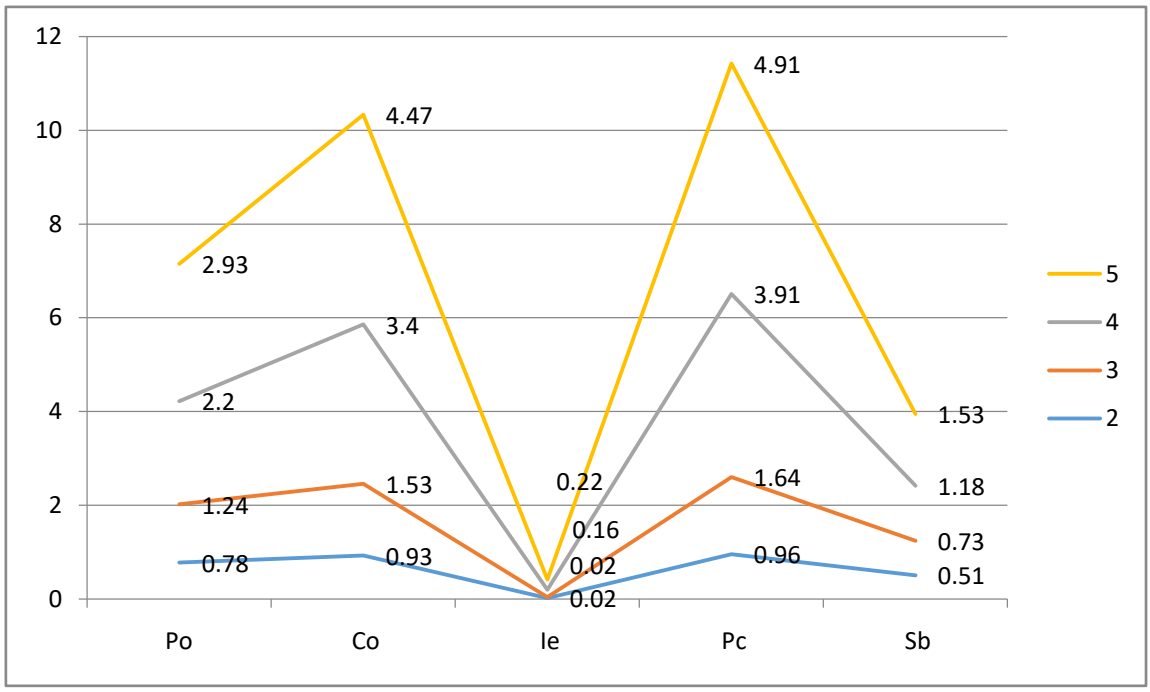

Figure 4.1 ***Mean Frequencies of the Numbers of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners***

As shown in Figure 4.1, the higher ratings the learners received, the more use of lexical chunks in each of the five categories. This result consists with the finding done by HE (2016), which found that the more advanced learners, the higher frequency and the more diversity of lexical chunks have been used.

Table 4.5

One Way ANOVA of the Numbers of Five Categories of Lexical Chunks

	Po	Co	Ie	Pc	Sb
df (df_{BW}, df_{WG})	(3,176)	(3,176)	(3,176)	(3,176)	(3,176)
F	21.59*	34.68*	3.66	34.57*	10.88*

Note: * Significant at $p < .01$

As shown in Table 4.5, the results of One-way ANOVA exhibit that significant differences existed in the mean frequencies of polywords (F=21.59, p<.01), collocations (F=34.68, p<.01), phrasal constraints (F=34.57, p<.01), and sentence builders (F=10.88, p<.01). No significant difference was found in the mean frequencies of institutionalized expressions (F=3.66, p>.01).

To further examine if any significant between-group difference existed among the four groups of EFL learners, post-hoc tests were performed on the data. Because the number of institutionalized expressions was much fewer than the number of other four categories, it was not analyzed with the tests. Table 4.6 shows the results of the post-hoc tests.

Table 4.6

Post-Hoc Tests of Mean Difference among Four Categories of Lexical Chunks

Paired Ratings	Po	Co	Pc	Sb
2--3	-.47	-.60	-.69	-.22
2--4	**-1.42***	**-2.47***	**-2.96***	**-.67***
2--5	**-2.16***	**-3.53***	**-3.96***	**-1.02***
3--4	**-.96***	**-1.87***	**-2.27***	**-.44**
3--5	**-1.69***	**-2.93***	**-3.27***	**-.80***
4--5	-.73	-1.07	-1.00	-.36

Note: * Significant at p< .01

The results indicate all of the paired comparison showed significant differences except 2 & 3-point groups and 4 & 5-point groups. The 2-point group did not differ from the 3-point group in their use of lexical chunks in all categories. Similarly, the 4-point group did not differ from the 5-point group in their use of all categories. What is noteworthy is that the passing groups (4-point and 5-point) used significantly more lexical chunks than the failing groups (2-point and 3-point). The results suggest that the holistic score of 4-point is a predictor of EFL learners' use of lexical chunks as well as a passing threshold for the GEPT Intermediate level writing test.

In terms of the numbers of different categories of lexical chunks used by EFL learners receiving different ratings, the learners tend to use phrasal constraints the most and use institutionalized expressions the least. Accordingly, the five categories of lexical chunks are not equally complicated for them. All four groups use phrasal constrains the most, which means that it is not a difficult task for the EFL learners to bear phrasal constraints in mind and to apply them in sentences. Simple words or phrases for the slots can be efficient to organize sentences. For example: a lot/kind/part of, in the future/world/past, a week/month/year ago. Even in the 2-point group, the most frequent use of phrasal constraints among the five categories can prove the explanation. With regard to institutionalized expressions, each rating group used them least. On the one

hand, the result echoes the previous literature review. Lewis (1993) proposed institutionalized expressions mainly occur in spoken language; and on the other hand, institutionalized expressions appeared in writing were often proverbs. For example, *Practice makes perfect* appeared several times in the passing groups. Nevertheless, it is not easy for the EFL learners in either passing groups or failing groups to memorize and produce an amount of institutionalized expressions in writing. Hence, the mean of the number of institutionalized expressions is quite low.

Subsequently, the results of One-way ANOVA show that there is a significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of different categories of lexical chunks. Furthermore, a significant difference exists between the passing group (4 & 5-point) and the failing group (2 & 3-point). Hence, the use of different types of lexical chunks is an indicator for passing or failing the GEPT writing test. The finding is similar to the one of Hsu's study (2007). Hsu found that the variety (type) of lexical collocations in comparison with the frequency of lexical collocations is a better indicator for the students' writing scores.

Lastly, notwithstanding a significant difference exists between 3-point group (failing group) and 4-point group (passing group) in their use of polywords, collocations, and phrasal constraints except sentence builders. This is probably because the mean

frequencies of sentence builders used in 3-point group and 4-point group are close. Moreover, the EFL learners in the both groups the most frequently used *I think/wish/consider/believe/hope/prefer/expect/know that* as sentence builders to express opinions or ideas. Apparently, the sentence builder, *I + verb + that …*, is a simple lexical chunk to produce the writing, especially when EFL learners do not have a variety of sentence builders in expressing their viewpoints. Similarly, Zonghui (2016) suggested that *I think*, as an opinion-presented lexical chunk (OPLC), is most repetitively used by Chinese students. When students don't have variety in opinion expressing, they typically overuse a certain phrases or a limited number of chunk, it shows that learners lack enough phrasal repertoires to employ in a native-like manner.

CHAPTER 5

CONCLUSION

This chapter concludes the present study on EFL learners' use of lexical chunks. It comprises the following sections: summary of the major findings, pedagogical implications, limitation of the study, and suggestions for future studies.

Summary of the Major Findings

The present study examined and compared the use of lexical chunks by Taiwanese EFL learners who received different ratings on the GEPT writing test. One hundred and eighty writing samples of the GEPT intermediate level with 45 samples from each rating group were analyzed and coded; their lexical chunks were identified and classified into five categories. One-way ANOVA was carried out to investigate the difference among Taiwanese EFL learners who received different ratings in terms of the total numbers of lexical chunks and five rounds of One-way ANOVA as well as Post-hoc comparisons were performed on the numbers of five categories of lexical chunks. The major findings are summarized in the following sections.

Difference in the Total Numbers of Lexical Chunks across Four Ratings of Writing by EFL Learners

The descriptive results of the total number of lexical chunks among different ratings show that the better learners perform on the GEPT writing test, the more lexical

chunks they used. Moreover, the results of One-way ANOVA show a significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of the total number of lexical chunks. Therefore, a post-hoc analysis was conducted to compare all paired means. The study found a significant difference between the passing groups (4 & 5-point) and the failing groups (2 & 3-point).

Difference in the Numbers of Five Categories of Lexical Chunks across Four Ratings of Writing by EFL Learners

The descriptive results of each category of lexical chunks reveal that the higher ratings, the more polywords, collocations, institutionalized expressions, phrasal constraints and sentence builders were used. By looking at each category of lexical chunks, the EFL learners in the 5-point group used more lexical chunks than other groups. Those in the 2-point group, on the other hand, used the fewest lexical chunks. Among different ratings, phrasal constraints were the most frequently used category of lexical chunks, followed by collocations, polywords, and sentence builders. Institutionalized expressions were the least used one of the five categories.

Moreover, the results of One-way ANOVA clearly indicate that there is a significant difference among Taiwanese EFL learners who received different ratings on the GEPT writing test in terms of the number of different categories of lexical chunks.

That is, the more advanced learners, the higher frequency and the more diversity of lexical chunks have been employed in writing.

What's more, a significant difference was found between the passing groups (4-point and 5-point) and the failing groups (2-point and 3 point), but within the passing groups or within the failing groups. The results suggest that the category of lexical chunks is an indicator for passing or failing in the GEPT writing test.

Pedagogical Implications

Based on the major findings of the present study, some pedagogical implications are drawn for EFL instruction in Taiwan. Firstly, EFL teachers in Taiwan are suggested to emphasize not only individual vocabulary and grammatical rules, but also lexical chunks. Lexical chunks should play a more essential role in language production and proficiency. Teachers can raise learners' awareness of lexical chunks by guiding students to highlight the lexical chunks in the given context and to identify each type of lexical chunks in both classroom activities and after-class exercises. By identifying chunks, students should be able to discriminate lexical chunks from random word combinations by themselves.

Secondly, teachers are suggested to teach fundamental lexical chunks first in order to facilitate the acquisition of native-like proficiency. Initial instruction may focus on

relatively fixed chunks that occur frequently including phrasal constraints, collocations, and polywords, and then pay attention to more variable chunks. Besides, lexical chunks may be learned better by teaching lexical chunks within a topic framework; that is, they are presented systematically in rich context rather than randomly. The method may assist EFL learners to keep lexical chunks in mind since it is effortless for learners to retrieve relevant lexical chunks when they meet similar topics again. With the development of Internet technology, using corpora in class is prevalent. On the one hand, it may help teachers choose appropriate corpora for students with different levels, needs, and interests, and on the other hand, it may encourage learners to find out the usage of lexical chunks by themselves.

Furthermore, teachers can introduce and emphasize the functions of lexical chunks in writing. For example, sentence builders provide the framework of the sentences. When introducing the chunks like *It is suggested that…*; *There is no doubt that…*, teachers can remind the learners that the functional effect of this kind of lexical chunks will not only signal the direction of the whole article, but also make the article more comprehensible. Even though institutionalized expressions are not needed much in writing, a proper one can make the writing native-like fluent. Hence, the learners may become aware of the importance of lexical chunks and also have a deep impression of

lexical chunks.

In a word, the most important of all is to raise learners' awareness of the importance of lexical chunks first, and then learn to use them appropriately. It is believed that lexical chunks will be particularly significant to L2 learners' language acquisition and production.

Limitations of the Study

Although the study has shed light into the relationship between the use of lexical chunks and writing performance, there are still two major limitations. Firstly, the present study only analyzed 180 samples, which limit the representation of the samples, and the finding may not easily generalized beyond the samples of the study.

Second, in the coding procedure of the present study, based on the representative corpus and dictionaries, lexical chunks in the writing samples were identified manually by the coders to avoid the personal subjectivity. The more coders involved in the identification of lexical chunks, the more reliable the result will be.

Suggestions for Future Research

Research on lexical chunks has been conducted in the past decades, but more studies can be done to sharpen our understanding of lexical chunks. As the present study

explored Taiwanese EFL learners' use of lexical chunks in relation to their writing performances, some suggestions can be made as follows. First, future studies are suggested to have a larger sample size to investigate the relationship between the use of lexical chunks and writing performances. Thus, Taiwanese EFL learners' use of lexical chunks in writing could be further examined. Moreover, the length of writing can also be considered to explore whether the longer the EFL learners write, the more lexical chunks they would produce.

Second, future research could investigate the correlation between the use and even knowledge of lexical chunks and other language skills (listening, speaking, and reading). Furthermore, other writing samples from other levels of GEPT such as elementary level, high-intermediate level, advanced, and superior level, and even the English writing samples of College Entrance Examination may obtain revealing findings.

Finally, it is suggested that future studies examine Taiwanese EFL learners' use of other classifications of lexical chunks in the writing samples. In addition to the Nattinger & DeCarrico's and Lewis' classifications used in the current study, there are other classifications, such as the functional classification including referential, stance, and discourse organizing (Biber et al, 2004). It could be a new direction to keep on studying whether there are some patterns of development in the EFL learners' use and

knowledge of lexical chunks especially related to the different categories.

REFERENCES

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

Astika, G. G. (1993). Analytical assessments of foreign students' writing. *RELC journal*, *24*(1), 61-70.

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

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers*, *26*, 181-190.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied linguistics*, *25*(3), 371-405.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E.(1999). *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language teaching research, 10*(3), 245-261.

Bygate, M. (1988). Units of oral expression and language learning in small group interaction. *Applied linguistics*, *9*(1), 59-82.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes*, *23*(4), 397-423.

Cowie, A. P. (1992). Multiword lexical units and communicative language teaching. In *Vocabulary and applied linguistics* (pp. 1-12). Palgrave Macmillan UK.

Cowie, A. P., & Howarth, P. (1996). Phraseological competence and written

proficiency. *British Studies In Applied Linguistics*, *11*, 80-93.

Ellis, N. C., Simpson-vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375-396.

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *Tesol Quarterly*, *28*(2), 414-420.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, *32*(4), 365-387.

Forsberg, F. (2010). Using conventional sequences in L2 French. *IRAL-International Review of Applied Linguistics in Language Teaching*, *48*(1), 25-51.

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. *Researching pedagogic tasks: Second language learning, teaching, and testing*, 75-93.

Graham, S. (2006). Writing. *Handbook of educational psychology*, 457–478.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. *Phraseology: Theory, analysis, and applications*, *145-160*.

Hakuta, K. (1976). Becoming bilingual: A case study of a Japanese child learning English. *Language learning*, *26*(2), 321-351.

Haswell, R. H. (1991). *Gaining ground in college writing: Tales of development and interpretation*. Southern Methodist Univ Pr.

Hausmann, F.J. (1999). Collocations in monolingual and bilingual English dictionaries. In I. Vladimir & D. Kalogjera (Eds.), Languages in contact and contrast: Essays in contact linguistics (pp. 225–236). Berlin: Mouton de Gruyter.

Hsu, J. Y. (2007). Lexical Collocations and Their Impact on the Online Writing of

Taiwanese College English Majors and Non-English Majors. *Online Submission*.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of second language writing*, *16*(3), 148-164.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation.*English for specific purposes*, *27*(1), 4-21.

Johnson, M. D., Acevedo, A., & Mercado, L. (2016). Vocabulary Knowledge and Vocabulary Use in Second Language Writing. *TESOL Journal*, *7*(3), 700-715.

Kimmes, A-M. (2004). An investigation of the usage and collocability of English verbs of thinking based on the online edition of the New York Times. Unpublished masters thesis, Southern Connecticut State University, New Haven, Connecticut, USA.

Kuiper, K., & Haggo, D. (1984). Livestock auctions, oral poetry, and ordinary language. *Language in society*, *13*(02), 205-234.

Laufer, B. (1992). How much lexis is necessary for reading comprehension?. In *Vocabulary and applied linguistics* (pp. 126-132). Palgrave Macmillan UK.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language learning*, *54*(3), 399-436.

Lewis, M. (1993). *The lexical approach*. Hove: Language Teaching Publications.

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

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, *18*(2), 85-102.

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

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. *Language and communication*, *191*, 225.

Peters, A. M. (1983). *The units of language acquisition* (Vol. 198, pp. 1-41). Cambridge: Cambridge University Press.

Qader, H. B. A. (2016). The Effect of Lexical Chunks on Kurdish EFL Learners' Writing Skill. *Education*, *6*(4), 101-106.

Richards, J. C.,& Rodgers, T. S. (2001). Approaches and methods in language teaching (2nd ed.). Cambridge: Cambridge University Press.

Schmitt, N. (2000). Key concepts in ELT. *ELT journal*, *54*(4), 400-401.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action. *Formulaic sequences: Acquisition, processing and use*, 1-22.

Snellings, P., Van Gelderen, A., & De Glopper, K. (2004). The effect of enhanced lexical retrieval on second language writing: A classroom experiment. *Applied Psycholinguistics*, *25*(2), 175-200.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for academic purposes*, *12*(3), 214-225.

Wood, D. (2002). Formulaic Language in Acquisition and Production: Implications forb Teaching. canada. *TESL Canada Journal*, *20*.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied linguistics*, *21*(4),463-489.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, *20*(1), 1-28.

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

Zhu, L. (2013). Analysis of prefabricated chunks used by second language learners of different levels. *Theory and Practice in Language Studies*, *3*(9), 1667.

Zonghui, H. E. (2016). A Contrastive Study on the Use of Lexical Chunk Among Chinese Learners of Different Proficiency Levels. *Studies in Literature and Language*, *12*(3), 64-70.

APPENDICES

APPENDIX A

Three Writing Topics and Prompts for the Current Study

試卷號碼：_____
試卷別：IW-0802

全民英語能力分級檢定測驗
GENERAL ENGLISH PROFICIENCY TEST

中級寫作能力測驗

二、英文作文（60%）

說明：請依下面所提供的文字提示寫一篇英文作文，長度約 120 字（8 至 12 個句子）。作文可以是一個完整的段落，也可以分段。（評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。）

提示：很多人都有崇拜的偶像 (idol)，除了影視名人、運動員或作家等外，也有可能是身邊的人。請寫一篇文章

(1) 描述你現在或以前所喜歡的偶像；

(2) 並說明你喜歡的理由。

卷號碼：_____
試卷別：IW-0862

全民英語能力分級檢定測驗
GENERAL ENGLISH PROFICIENCY TEST

中級寫作能力測驗

二、英文作文（60%）

說明：請依下面所提供的文字提示寫一篇英文作文，長度約 120 字（8 至 12 個句子）。作文可以是一個完整的段落，也可以分段。（評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。）

提示：歷史上有許多發明改變了人類的生活。請寫一篇文章

(1) 描述未來你最希望看到的一項發明；

(2) 說明這項新發明的重要性。

試卷號碼：_____

試卷別：IW-0863

全民英語能力分級檢定測驗
GENERAL ENGLISH PROFICIENCY TEST

中級寫作能力測驗

二、英文作文（60%）

說明：請依下面所提供的文字提示寫一篇英文作文，長度約 120 字（8 至 12 個句子）。作文可以是一個完整的段落，也可以分段。（評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。）

提示：音樂是許多人生活中不可或缺的一部份。請寫一篇文章

(1) 說明音樂對你的重要性並陳述理由；

(2) 描述你自己最喜歡的音樂類型。

APPENDIX B

Coding Sheet

WCs	❶Po.	❷Co.	❸Ie.	❹Pc.	❺Sb.	Total
Errors						

WCs: word counts ❶Po: Polywords ❷Co: Collocations ❸Ie: Institutionalized expressions ❹Pc: Phrasal constraints ❺Sb: Sentence builders)

(I) Features and Examples of Different Types of Lexical Chunks

Classifications	length	variation	Examples
❶ Polywords **(Po)**	phrase level	fixed	pull one's leg, run into, turn down, get along with, after all, at any rate, at any time, on the other hand, to sum up,…
❷ Collocations **(Co)**	phrase level	fixed or variable (productive)	
① verb＋noun			go shopping, play basketball,…
② adjective＋noun			strong tea, dirty words, heavy rain,…
③ noun + verb naming an action			bomb explodes, bees buzz,…
④ adverb＋adjective			sound asleep, fully aware,…
⑤ verb + adverb			whisper softly, argue heatedly, …
⑥ noun + noun			dress code, entrance exam,…
❸ Institutionalized expressions **(Ie)**	sentence level	fixed	How are you? Seeing is believing. Birds of a feather flock together.
❹ Phrasal constraints **(Pc)**	phrase level	variable (with slots)	a ____ ago, the ___er the ____er, as _____ as, both _____ and _____, a_____ of ____,…
❺ Sentence builders **(Sb)**	sentence level	variable (with slots)	I think that…/ not only…, but also…/ firstly,…secondly,…finally,…

(II) Definitions of Different Types of Lexical Chunks

Polywords are fixed short phrases with no variability, and can be idioms (e.g., pull one's leg), phrasal verbs (e.g., run into, turn down, get along with), adverb phrases (e.g., after all, at any rate, at any time, and so on) and phrases functioning as transitions (e.g., on the other hand, to sum up). The meaning of the whole chunk may be apparent or opaque, and can or cannot be inferred from the meaning of the individual words.

Collocations refer to a string of words that co-occur in a natural text with greater than random frequency. According to Benson et al. (1997), Hausmann (1999), and Kimmes (2004), there are six types of collocations, including "verb＋noun (e.g., go shopping), adjective＋noun (e.g., heavy rain), noun + verb naming an action (e.g., bees buzz), adverb＋adjective (e.g., sound asleep), verb + adverb (e.g., whisper softly), noun + noun (e.g., school uniform)".

Institutionalized expressions are sentence-length and invariable chunks (e.g., How are you? Practice makes perfect. Not yet. See you later). They are proverbs, aphorisms, formulas for social interaction that are stored as complete units in the mental lexicon.

Phrasal constraints are short-to-medium-length phrases and allow variation of to-be-filled slots for the lexical elements (e.g., as ____ as, a ___ ago, the ___er the ___er), including noun phrases (e.g., a piece of, a lot of), adjective phrases (e.g., afraid of, good at, famous for), and prepositional phrases (e.g., in the future, in history, around the world), etc..

Sentence builders provide the framework for whole sentences to be constructed (e.g., I think that…/ Firstly,…Secondly, … Finally,…/ It is suggested that…/ My point is that…). They allow variation of to-be-filled slots for the language users to express ideas.

APPENDIX C

Coding Samples

卷號碼：_____
試卷別：IW-0862

全民英語能力分級檢定測驗
GENERAL ENGLISH PROFICIENCY TEST

二、英文作文（60%）

說明：請依下面所提供的文字提示寫一篇英文作文，長度約 120 字（8 至 12 個句子）。作文可以是一個完整的段落，也可以分段。（評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。）

提示：歷史上有許多發明改變了人類的生活。請寫一篇文章
(1) 描述未來你最希望看到的一項發明；
(2) 說明這項新發明的重要性。

IW0862-5
【2266】

Human beings❷ always make attempts❷ to create useful inventions❷ with a view to❶ improving our daily life❷. From my perspective❹, I hope❺ we can have "Time machine❷" in the future❹. We always see this invention in movies or cartoons. It has the ability to bring us back to❹ ancient times❷. Maybe it doesn't have attractive appearance❷, but it has an unbelievable power.

If we have the "Time machine" in the future, we'll be capable of❹ changing the mistakes we did before. Furthermore, we can know our future through this invention. It can satisfy people because we get the power of altering things. However, some people may use this machine to commit crimes❷ secretly. In order to❶ prevent these people from❹ using this invention in a bad way❹, we must discuss strict laws❷ and put more emphasis on❹ ethics. As a result❶, we'll be truly no more❷ confined to❹ the time,

since we combine technology and morality successfully❷.

WCs	❶Po.	❷Co.	❸Ie.	❹Pc.	❺Sb.	Total
157	3	10	0	8	1	22
Errors	----					
0						

IW0862-4

【2084】

Some people don't like to have new invention❷. They think❺ new invention makes people become lazier. However, in the future❹, I want to have a new invention-a time machine❷, because people don't have enough time!

A time machine can solve many problems❷ in the world❹. For example❶, students usually have lots of ❹homework, tests, and classes. The time machine can make time❷ for them to study courses❷, prepare tests.. etc. Businessman always have to do lots of work a day. They don't have enough time to work, and rest. The time machine can make every thing go slower. That's why I wish❺ there will be a time machine someday.

WCs	❶Po.	❷Co.	❸Ie.	❹Pc.	❺Sb.	Total
109	1	5	0	3	2	11
Errors	prepare^ “for” tests					
1						

IW0862-3

【2199】

Many invention have changed people's life in history❹. MRT makes convenient. Computer lets people get more information. But the argument and fighting still happen.

There is no machine to make the world peace. Because of that, so I hope someone will invent a machine❷ to make the world peaceful.

There are more than one thousand refugees in Africa. There is poverty instead of peace. In the 2008, China happened a disaster-earthquake. Many children haven't have their parents ever. If we have a machine to make the world peaceful. These tragedy won't happen ever. I want to see a wonderful world❷, a peaceful world.

WCs	❶Po.	❷Co.	❸Ie.	❹Pc.	❺Sb.	Total
103	0	2	0	1	0	3
Errors	Because of that,…so I hope…					
2	There is poverty instead of peace.(?)					

IW0862-2
【2217】

In the future❹, I wanna to create the civil engineering robot. This one can do a lot of❹ things that can't be done now.

Now days, there are earthquakes and typhoons in Taiwan. And too many buildings are broken by natural power. So the civil engineering robot is the one that can solve all the problems❷.

The most important thing is that the robot don't care of safety. Because it is no life.

WCs	❶Po.	❷Co.	❸Ie.	❹Pc.	❺Sb.	Total
73	0	1	0	2	0	3
Errors	The most important thing is that the robot don't care of safety.(→doesn't					
1	care about its security.)					

APPENDIX D

Intercoder Agreement

Samples	(A) Total Number of Corder A	(B) Total Number of Corder B	(C) Number of codings that two coders have discrepancies	(D) Total number of codings (A+B=D)	(E) Number of codings that two coders both agree (D-C=E)
2-point	145	100	45	245	200
3-point	241	152	89	393	304
4-point	483	386	97	869	772
5-point	641	572	69	1213	1144
Sum	1510	1210	300	2720	2420

Intercoder Agreement: 89% (Brown, 2001)

=Number of codings that two coders both agree (E) / Total number of codings (D) ×100

(2420÷2720＝0.88970×100＝88.97%≒89%)

APPENDIX E

Two Writing Topics and Prompts (Mini Pilot Study)

全民英檢中級寫作能力測驗第二部分「英文作文」

例題：

例題 I

說明:請依下面所提供的文字提示寫一篇英文作文,長度約120字(8至12個句子)。作文可以是一個完整的段落，也可以分段。(評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。)

提示：一般來說，孩子表現良好時，父母通常會給孩子獎勵。請寫一篇文章說明
（1） 你表現好的時候，你的父母通常會用哪些方法獎勵你？你覺得這些方法有、適當嗎？
（2） 如果有一天你為人父母，你會用相同的獎勵方式嗎?

例題 II

說明: 請依以下提示寫一封英文 email，長度約 120 字 (8 至 12 個句子)。此信可以是一個完整的段落，也可以分段。(評分重點包括內容、組織、文法、用字遣詞、標點符號、大小寫。)

提示：你住在美國的筆友 Ted 最近剛轉學，正在煩惱新學校的交友問題。請寫一封英文 email 給他，信的內容必須包括以下兩點：
（1） 分享你適應新環境的經驗。
（2） 建議他幾個交新朋友的方法。

Email 的上下款不須寫在答案紙上

(Retrieved from: https://www.lttc.ntu.edu.tw/geptscoreremark/icomposition.pdf)

APPENDIX F

Intercoder Agreement (Mini Pilot Study)

Samples		(A) Total Number of Corder A	(B) Total Number of Corder B	(C) Number of codings that two coders have discrepancies	(D) Total number of codings (A+B=D)	(E) Number of codings that two coders both agree (D-C=E)
5-point	1	22	19	3	41	38
	2	17	17	2	34	32
4-point	3	18	18	2	36	34
	4	18	17	3	35	32
3-point	5	8	8	0	16	16
	6	7	6	3	13	10
2-point	7	6	5	1	11	10
Sum		96	90	14	186	172

Intercoder Agreement: 92% (Brown, 2001)

=Number of codings that two coders both agree (E) / Total number of codings (D) ×100

(172÷186＝0.9247×100＝92.47%≒92%)

APPENDIX G

Distribution of Five Categories of Lexical Chunks (Mini Pilot Study)

Samples		Coder A & Coder B					
		Polywords	Collocations	Institutionalized expressions	Phrasal constraints	Sentence builders	Total Number
5-point	1	5	11	0	3	3	22
	2	1	10	0	4	2	17
4-point	3	1	10	1	3	4	19
	4	2	12	0	2	3	19
Passing	↑	9	43	1	12	12	77
Passing	Mean	2.25	10.75	0.25	3.00	3.00	19.25
Failing	Mean	2.33	2.33	0.33	0.33	2.00	7.33
Failing	↓	7	7	1	1	6	22
3-point	5	1	5	0	0	2	8
	6	4	0	1	0	3	8
2-point	7	2	2	0	1	1	6