

ผลกระทบของการสอนการปรากฏร่วมจำเพาะของคำที่มีต่อการ
ใช้การปรากฏร่วมจำเพาะของคำของนักเรียนที่เรียน
ภาษาอังกฤษในฐานะภาษาต่างประเทศ

นางชุติง โฮ่ว

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรมหาบัณฑิต
สาขาวิชาภาษาอังกฤษศึกษา
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2555

**THE EFFECTS OF COLLOCATION INSTRUCTION ON
NON-ENGLISH MAJOR EFL STUDENTS'
USE OF COLLOCATIONS**

Quping Hou

**A Thesis Submitted in Partial Fulfillment of the Requirements for
the Degree of Master of Arts in English Language Studies**

Suranaree University of Technology

Academic Year 2012

**THE EFFECTS OF COLLOCATION INSTRUCTION
ON NON-ENGLISH MAJOR EFL STUDENTS'
USE OF COLLOCATIONS**

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee

(Assoc. Prof. Dr. Anchalee Wannaruk)

Chairperson

(Asst. Prof. Dr. Issra Pramoolsook)

Member (Thesis Advisor)

(Dr. Sirinthorn Seepho)

Member

(Dr. Jitpanat Suwanthep)

Member

(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs

(Dr. Peerasak Siriyothin)

Dean of Institute of Social Technology

ชูผิง โฮว้ : ผลกระทบของการสอนการปรากฏร่วมจำเพาะของคำที่มีต่อการใช้การปรากฏ
ร่วมจำเพาะของคำของนักเรียนที่เรียนภาษาอังกฤษในฐานะภาษาต่างประเทศ
(THE EFFECTS OF COLLOCATION INSTRUCTION ON NON-ENGLISH MAJOR
EFL STUDENTS' USE OF COLLOCATIONS) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์
ดร.อิสรา ประมูลสุข, 148 หน้า.

เรื่องยากประการหนึ่งที่ผู้เรียนภาษาต่างประเทศประสบบ่อยในการเขียน คือ การเลือกคำมา
ใช้ให้เหมาะสมเหมือนเจ้าของภาษา ในประเทศจีนแม้ว่าจะมีงานวิจัยและความรู้เกี่ยวกับเรื่องการ
เรียนการสอนคำศัพท์มากมายและหลากหลาย แต่การศึกษาเกี่ยวกับการปรากฏร่วมจำเพาะของ
คำศัพท์ยังเป็นที่ต้องการอยู่ ดังนั้น ในงานวิจัยชิ้นนี้ ในการศึกษาแรกผู้วิจัยตั้งใจที่จะระบุ
ข้อผิดพลาดเกี่ยวกับการปรากฏร่วมจำเพาะของคำศัพท์ในการเขียนของนักศึกษาชาวจีนที่ไม่ได้เอก
วิชาภาษาอังกฤษที่มหาวิทยาลัยไคลี่ ซึ่งมีวัตถุประสงค์ที่จะช่วยผู้วิจัยตัดสินใจว่าจะใช้ Corpus of
Contemporary American English (COCA) อย่างไร เพื่อเพิ่มพูนความตระหนักในเรื่องการปรากฏ
ร่วมจำเพาะของคำศัพท์ ในการวิจัยนี้ผู้วิจัยได้ขอให้นักศึกษาที่ไม่ได้เอกวิชาภาษาอังกฤษ จำนวน
150 คน เขียนชิ้นงานคนละ 2 ชิ้น ซึ่งทำให้ได้ชุดข้อมูลที่ประกอบไปด้วยความเรียง 300 ชิ้น จากนั้น
จึงทำการหาและระบุข้อผิดพลาดการปรากฏร่วมจำเพาะของคำศัพท์ โดยใช้ COCA เป็นชุดข้อมูล
อ้างอิง ผลการวิจัยพบว่า ข้อผิดพลาดการปรากฏร่วมจำเพาะของคำที่พบบ่อยที่สุด คือ การปรากฏ
ร่วมจำเพาะที่มีคำกริยาเป็นแกนคำ และที่รองลงมาคือ การปรากฏร่วมจำเพาะที่มีคำคุณศัพท์เป็น
แกนคำ นอกจากนี้ ยังพบว่านักศึกษาใช้คำเพื่อบอกปริมาณผิดพลาด และมักจะเลือกใช้คำกริยา
พิเศษกับคำกริยาผิดมากกว่า เมื่อเทียบกับการเลือกใช้คำกริยาพิเศษกับคำคุณศัพท์

นอกจากนั้น ในการศึกษาที่ 2 นักวิจัยได้ทำการศึกษาถึงการทดลองเวลาแปลสัปดาห์ เพื่อดู
ถึงผลของการใช้ COCA ในการเพิ่มพูนความตระหนักในเรื่องการปรากฏร่วมจำเพาะของคำศัพท์
โดยในสัปดาห์แรกเป็นการทดสอบก่อนเรียนและการเขียนชิ้นงานก่อนเรียน หกอาทิตย์หลังจากนั้น
เป็นการเรียนการสอน และในอาทิตย์สุดท้ายเป็นการทดสอบหลังเรียน การเขียนชิ้นงานหลังเรียน
และทำแบบสอบถาม ผลการวิจัยพบว่า ผู้เรียนมีพัฒนาการที่ดีขึ้นอย่างเห็น ได้ชัด ซึ่งเห็นได้จากการ
ทดสอบหลังเรียน นักศึกษาใช้การปรากฏร่วมจำเพาะของคำได้ถูกมากขึ้น และใช้ผิคน้อยลง แต่ผล
จากการเขียนชิ้นงานก่อนและหลังเรียนไม่เห็นผลชัดเจนมากนัก นอกจากนี้ จากแบบสอบถาม
นักศึกษาระบุว่ามีทัศนคติที่ดีต่อการ ใช้ COCA ในการเรียนรู้เรื่องการปรากฏร่วมจำเพาะของคำ

สาขาวิชาภาษาต่างประเทศ
ปีการศึกษา 2555

ลายมือชื่อนักศึกษา _____
ลายมือชื่ออาจารย์ที่ปรึกษา _____

QUPING HOU : THE EFFECTS OF COLLOCATION INSTRUCTION ON
NON-ENGLISH MAJOR EFL STUDENTS' USE OF COLLOCATIONS.

THESIS ADVISOR : ASST. PROF. ISSRA PRAMOOLSOOK, Ph.D.,

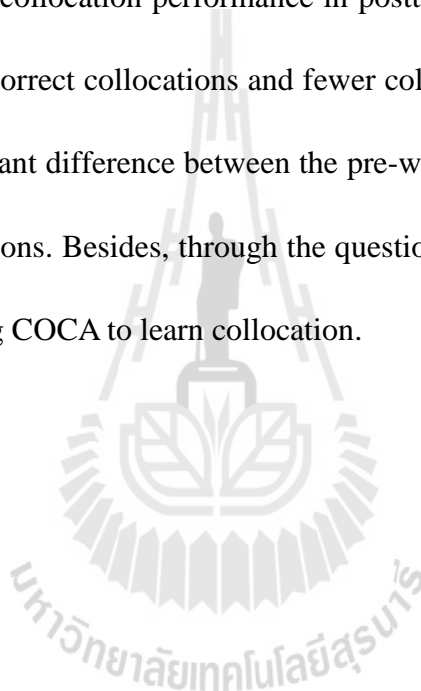
148 PP.

NON-ENGLISH MAJORS/LEXICAL COLLOCATION/COCA/ERROR

ANALYSIS/COLLOCATION AWARENESS

One of the difficulties EFL learners frequently experience in writing is the choice of words to achieve native-like competency. In China, though research reports and literature on the teaching and learning of vocabulary are very extensive, studies on collocations are still in need. Therefore, in this piece of research, the first study aimed to identify the lexical collocation errors made by Non-English Majors (NEMs) in Kaili University (KU) in their writing, which was intended to help the researcher to decide how to utilize Corpus of Contemporary American English (COCA) to raise their collocation awareness. Two writing tasks were administered to 150 NEMs in KU, resulting in a corpus of 300 essays. Lexical collocation errors in the texts were identified by two raters using COCA as a reference corpus. The results revealed that the most frequent collocation errors were collocations with verbs as nodes and the second most ones were collocations with adjectives as nodes. Misuses of quantifiers were also found in the corpus. Moreover, the students had more trouble in choosing an adverb to go with a verb than choosing an adverb to go with an adjective.

Then, an eight-week quasi-experimental study was employed in the second study to investigate the effects of utilizing COCA on raising learners' awareness concerning lexical collocations. This study included the first week for pre-test and pre-writing, six weeks of instructional treatment, and final week for posttest, post writing, and questionnaire. The findings showed that the participants achieved a significant improvement on their collocation performance in posttest after the 6-week treatment. They produced more correct collocations and fewer collocation errors in post writing, but there is no significant difference between the pre-writing and post writing on their use of lexical collocations. Besides, through the questionnaire they expressed positive attitudes towards using COCA to learn collocation.



School of Foreign Languages

Student's Signature _____

Academic Year 2012

Advisor's Signature _____

ACKNOWLEDGEMENT

The completion of this research work should be attributed to the help and care I have got from so many experts and friends. I am very grateful to all of them for what they have blessed me.

First of all, I would like to express my deepest gratitude to Dr. Issra Pramoolsook, my supervisor, who supported and encouraged me generously throughout this study. Without his expertise, patience, and guidance, it would not have been possible for me to complete this work.

I would like to express my gratitude to the advisory committee members Assoc. Prof. Dr. Anchalee Wannaruk, Dr. Jitpanat Suwanthep, and Dr. Sirinthorn Seepho for their recommendations, suggestions, valuable comments, advice, and warm encouragement.

My sincere thanks also go to the course instructors at School of Foreign Languages, Suranaree University of Technology, for their lectures in MA courses, which have always been a source of academic guidance for me over the writing of this thesis.

My appreciation is also extended to my colleagues in Kaili University, Guizhou Province, China for their kind support and warm encouragement. I would be grateful to those who participated in the pilot study and the main study for their help with my

data collection and conducting the research. Without them, my research would not have been possible.

And my heartfelt thanks go to my family and my beloved friends for their love, understanding, encouragement, and support throughout the course of my study.

Quping Hou

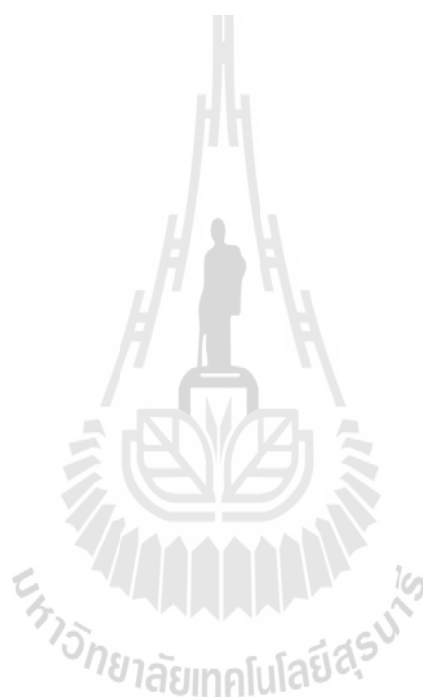


TABLE OF CONTENTS

	Page
ABSTRACT (THAI)	I
ABSTRACT (ENGLISH).....	II
ACKNOWLEDGEMENTS	IV
TABLE OF CONTENT	VI
LIST OF TABLES.....	X
LIST OF FIGURES.....	XII
LIST OF ABBREVIATION.....	XIII
CHAPTER	
1. INTRODUCTION.....	1
1.1 Research Background.....	1
1.2 Problem Statement.....	5
1.3 Rationale of the study.....	6
1.4 Purposes of the Study.....	7
1.5 Research Questions.....	8
1.6 Significance of the Study.....	8
1.7 Definitions of Key Terms.....	10
1.8 The Overall Structure of the Study.....	11

TABLE OF CONTENTS (Continued)

	Page
2. LITERATURE REVIEW	13
2.1 Collocation.....	13
2.1.1 Definitions of Collocation.....	13
2.1.2 Collocation Competence.....	21
2.1.3 Classification of Collocations	23
2.1.3.1 Collocation as a continuum.....	24
2.1.3.2 Collocation restriction.....	24
2.1.3.3 Collocation strength.....	25
2.1.3.4 Word frequency of collocations.....	26
2.1.3.5 Characteristics and functions of collocations.....	27
2.2 Error Analysis.....	31
2.3 Sources of Collocation Errors.....	35
2.4 Previous Studies on Teaching English Collocations in EFL Contexts.....	37
2.5 Teaching of Collocations.....	44
3. RESEARCH THODOLOGY	49
3.1 Research Design.....	49
3.1.1 Phase I Error analysis.....	49
3.1.2 Phase II A quasi-experimental study.....	50

TABLE OF CONTENTS (Continued)

	Page
3.2 Participants and Data Collection.....	53
3.2.1 Participants of the Study.....	53
3.2.2 Data Collection.....	53
3.2.2.1 Data from Writing Tasks & Error Analysis.....	53
3.2.2.2 Data from Pre-test & Post-test.....	58
3.2.2.3 Data from Instructional Treatment and Questionnaire.....	63
3.2.2.4 Data from Interview.....	67
3.3 Data Analysis.....	69
3.3.1 Qualitative Analysis.....	69
3.3.2 Quantitative Analysis.....	69
3.4 Pilot Study.....	70
3.4.1 Participants and Procedures.....	60
3.4.2 Results of the Pilot Study.....	71
3.4.2.1 Results of Error Analysis.....	71
3.4.2.2 Piloting Instructional Treatment.....	79
3.4.2.3 Comparison between the Collocation Pre- and Post test.....	80
3.4.2.4 Results of the Questionnaire.....	82
3.5 Conclusion and Implications for the Main Study.....	85
3.6 Summary of Chapter 3.....	86

TABLE OF CONTENTS (Continued)

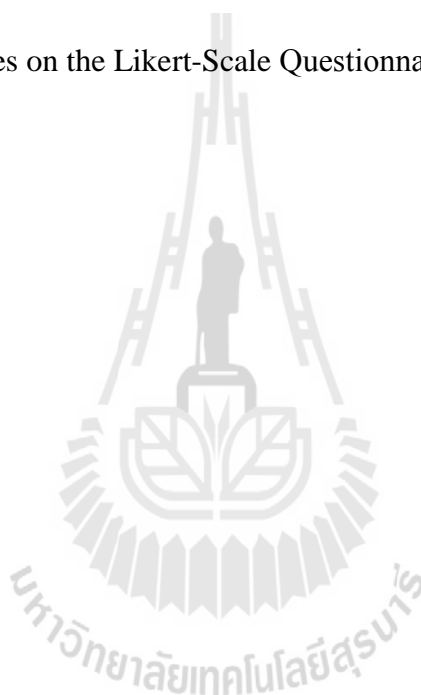
	Page
4. RESULTS AND DISCUSSION	87
4.1 Typical Lecical Collocation Errors in the Learner Corpus.....	87
4.2 Discussion.....	92
4.3 Effects of Collocation Instruction.....	96
4.3.1 Participants’ Performance on Pre- and Post Writing	97
4.3.2 Participants’ Performance on Pre- and Post Test.....	101
4.3.3 Psrticipants’ Attitudes as Revealed by the Questionnaire.....	101
4.4 Discussion.....	104
4.5 Summary of Chapter 4.....	106
5. CONCLUSION AND IMPLICATIONS	107
5.1 Conclusion of the Study.....	107
5.2 Pedagogical Implicantions.....	109
5.3 Limitation of the Study.....	111
5.4 Recommendations for Further Studies.....	113
REFERENCES	115
APPENDICES	126
CURRICULUM VITAE	148

LIST OF TABLES

Tables	Page
2.1 Grammatical Collocation Types.....	28
2.2 Lexical Collocation Types.....	29
2.3 Types of Vocabulary Learning Strategies (Nation, 2001).....	46
3.1 Summary of Instruments or Means to Answer Research Questions.....	51
3.2 Lexical Collocation Error Domains and Categories.....	56
3.3 Examples Extracted from COCA.....	57
3.4 The Distribution of Collocation Types in Collocation Test.....	60
3.5 The Number of Lexical Collocations Selected from CET Papers.....	62
3.6 Treatment Arrangement in the Present Study.....	66
3.7 Number and Percentage of Correct Collocation and Collocation Errors in Learner Corpus.....	73
3.8 Distribution of Lexical Collocation Error Domains (Pilot Study).....	76
3.9 Students' Responses from Student on the Likert-Scale Questionnaires (N=10)..	83
4.1 Correct and Incorrect Collocations in the two Writing Tasks.....	89
4.2 Distribution of Lexical Collocation Error Domains (Main Study).....	91
4.3 Comparison of Collocation Errors between Pre- and Post Writing.....	98
4.4 Correct and Incorrect Collocations in Pre- and Post Writing.....	99

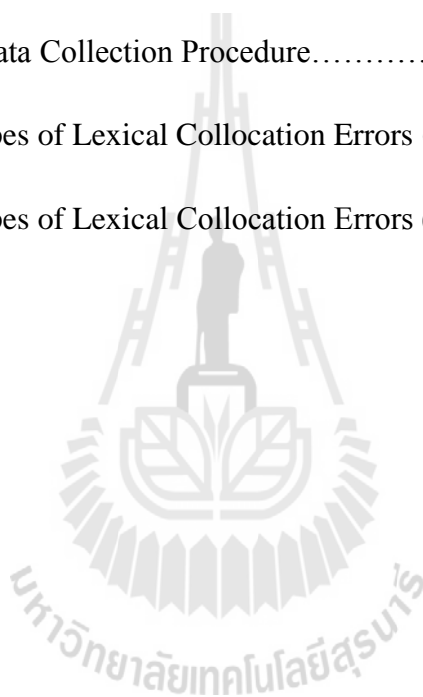
LIST OF TABLES (Continued)

Table	Page
4.5 Distribution of Lexical Collocation Error Domains in Pre- and Post Writing.....	100
4.6 Comparison of Collocation Test Scores between Pre- and Post Test.....	101
4.7 Students' Responses on the Likert-Scale Questionnaires (N=39).....	102



LIST OF FIGURES

Figure	Page
2.1 Arbitrariness of Collocation.....	19
2.2 Collocation Continuum	24
3.1 An Overview of Data Collection Procedure.....	52
3.2 Distribution of Types of Lexical Collocation Errors (Pilot Study)	74
4.1 Distribution of Types of Lexical Collocation Errors (Main Study).....	90



LIST OF ABBREVIATION

CAH	=	Contrastive Analysis Hypothesis
CET	=	College English Test
CET-SET	=	College English Test-Spoken English Test
COCA	=	The Corpus of Contemporary American English
CG	=	Control Group
EG	=	Experimental Group
GWDS	=	The Great Western Development Strategy
IOC	=	Item-Objective Congruence Index
KU	=	Kaili University
NEMs	=	Non-English majors
NCETC	=	The National College English Testing Committee
QMDAP	=	Qiandongnan Miao and Dong Autonomous Prefecture
VLT	=	The Vocabulary Levels Test

CHAPTER 1

INTRODUCTION

This chapter gives an introduction to the study which focuses on non-English major students' collocation errors at Kaili University in South West China. It covers the background, the problems, and the purposes of the study. Research questions and the significance of this study are also provided. It ends with an overall structure of the thesis.

1.1 Research Background

Nowadays, English has been used as an international language. With the globalization and the fast development of economy, English plays more and more indispensable roles in China. As in many countries, English study in China is inspired by not only the desire to study abroad, but also a need to improve skills and find good jobs. Western teaching methods and technology have been introduced into Chinese schools. Traditional emphasis on obtaining a huge vocabulary through recitation is gradually replaced by a more integrated approach to English writing, listening, reading and speaking. The Chinese Ministry of Education requires that language teaching in China in the 21st century should train students to be competent enough to use English for communication with the outside world.

Many universities in China have English majors and hire foreign teachers to help students with their written and spoken communicative skills. Even non-English majors can often have the chance to receive help from these foreign instructors. Not only do most schools and universities offer English courses in China, but also there are thousands of training centers that specialize in teaching English to children and adults. Adults who want to adapt themselves to the need of society are eager to learn and master a foreign language, especially English. Learning English in China is a prerequisite for any students. Chinese students who want to use English to further their career will have to do more than just becoming familiar with the language.

College English is an important and compulsory course for students of non-English majors (NEMs) in China. The objective of College English is to develop students' ability to use English in a well-rounded way so that in their future studies and careers as well as social interactions, they will be able to communicate effectively in both oral and written forms. As English study booms, many Chinese college students are taking College English Test (CET) to examine their English level. The CET is a national English language test for university students in China who are not English majors. It is a large-scale and standardized test designed and developed by the National College English Testing Committee (NCETC) which is directed and supervised by the Department of Higher Education of the Ministry of Education. Many employers in China prefer applicants with CET certification. Graduates are able to get good jobs only when they can pass the CET with high scores.

CET Band-4 and CET Band-6 are designed for college students who have completed the corresponding English courses and are aimed at precisely measuring college students' comprehensive command of English. Thus, CET plays an active role in realizing the objectives of college English teaching. According to the Curriculum Requirements, English teaching has three stages at the college level: basic, intermediate and advanced stages. CET Band-4 is designed for college students who have completed the basic stage of College English learning, while CET Band-6 is for college students who have completed the intermediate stage of College English learning. CET-SET (College English Test-Spoken English Test) is designed for college students, at an advanced stage, who have taken CET Band-4 or CET Band-6 and obtained the required scores (NCETC, n.d.). The CET test is held nationally twice a year in June and December and includes listening, reading and writing sections. The spoken test is optional and is employed to examine students' oral communication ability on familiar topics. Since the first CET Band-4 was held in 1987 and the first CET Band-6 in 1989, the number of students taking the national English test has been stably increasing every year. Before the 2005 reform, the maximum possible score of the test was 100 points. In 2005, several changes were made, including re-ordering and re-organizing sections of the test paper. More listening and speaking tasks have been involved, new sections of "Skimming and Scanning" and "Translation" have been introduced. The grading system has been modified (scores are graded on a curved scale, so that the highest score is 710 while the lowest is 290). Participants have been given detailed score reports without passing grades and qualification certificates.

Kaili University (KU) is a local public university located in the southeast of Guizhou Province, China. It is the largest higher education institution in Qiandongnan Miao and Dong Autonomous Prefecture (QMDAP) with a student population of over 10,000, among whom 6733 are NEMs, 955 are English majors, and 3168 are adult students. In 1999, the Chinese central government made an important decision to implement the Great Western Development Strategy (GWDS). The GWDS is a policy adopted by the central government to boost the development of the less developed western regions, one of which is Guizhou Province. The GWDS needs more people who have high levels of English proficiency especially in education fields. In accordance with the demand of promoting education of GWDS, the main goal of Kaili University is to cultivate teachers for schools in QMDAP.

Most of the NEMs in KU cannot achieve the required level of college English education. Almost every student in KU has the experience to take CET but few of them can pass it. Before they enter university, they learn English for at least ten years but their English proficiency is still at a low level. Since they cannot get good opportunities for foreign language learning at high schools, English learning is usually one of the toughest tasks for these students in KU. Zheng (2000) states that many college students with low reading speed cannot understand the contents of what they have read. There exist some common characteristics in English study, which impede the improvement of the ethnic undergraduates' listening and speaking (Wang, 2010). Among the four basic English learning skills, writing is the weakest for them. Many

problems exist in students' compositions, and the learners usually do not know how to choose words to express themselves clearly (Wu, 2003).

1.2 Problem Statement

English writing has always been an essential issue in English teaching. Work would be easier for language teachers if the students do not have problems in writing. However, for NEMs in KU, even though they have learnt thousands of words by heart, it is hard to achieve native-like written communication. Written ability can reflect a person's productive English proficiency. For these NEMs, one of the obstacles to learn English is the lack of competence in using collocations. They often use unnatural English expressions that have right word items but improper collocations in their spoken and written communication (Wu, 2003). Collocational idiosyncrasy always seems to be a great difficulty for NEMs in KU to overcome and often leads to embarrassment in spoken and written communication. Before entering university, they have been studying English for at least ten years, and they consider that they have mastered grammatical rules and a certain number of words, but they still have difficulty in making themselves understood or their speech natural.

The following expressions are very common collocation errors found in NEMs' writing in KU:

1. They **study knowledge** (*acquire knowledge*) *from teachers*.
2. *Wasting time is fatal* **by my useless opinion** (*in my opinion*).
3. *After dinner,* people **eat tea** *in the yard* (*drink tea*).

1.3 Rationale of the Study

“Collocation runs through the whole of the English language. No piece of natural spoken or written English is totally free of collocation.” (Oxford Collocations Dictionary, 2002, p. vii). The use of collocations has been recognized as one of the ways that differentiate native speakers and non-native speakers. It has been observed that native speakers usually have their ability to dispose thousands of words (Kjellmer,1991; Howarth,1998). However, in the attempt to achieve the competence of native-speakers, foreign/second language learners have not got enough 'chunks' of language or collocations to draw on. A student could easily invent the cumbersome phrase “*Her disability is forever*” because they lack the “adj + noun” collocation: “*She has a permanent disability*”. Howarth (1998) points out that not knowing which words to go with which, and do not go with others is a major problem for learners, which is true for Chinese university English learners (Meng & Li, 2005).

Errors are inevitable in the process of foreign language learning. Errors can tell teachers how far the learners are from the teaching objectives. As for students, errors are indispensable, too. They can learn from making errors. Error analysis can keep teachers and students focused closely on specific language rather than viewing universal aspects of language (Xie & Jiang, 2007). Students’ essay writing often shows a serious lack of collocation competence (Gitsaki,1999; Chen, 2002; Ying & Hendricks, 2004; Meng & Li, 2005; Fan, 2009).

In 1994, Tim Johns made a ground-breaking effort to bring corpus use to the classroom, which is data-driven learning. The idea is to guide learners to discover features and patterns in the foreign language through using corpora. Seidlhofer (2000), Kennedy & Miceli (2001), Chambers & O'Sullivan (2004), and Bernardini (2004) are researchers in favor of this approach.

COCA is the largest and most recent corpus of English, and it can provide large amount of authentic information of English. It is a window to observe the usage and changes of English. Students can enhance their language sense, grammar, and lexical system through studying concordances in COCA. In the present study, through analyzing NEMs' lexical collocation errors in KU, the researcher focused on students' weaknesses in using lexical collocations, and wanted to raise students' lexical collocation awareness by utilizing COCA to search collocates to a given node, correct lexical collocation errors, and collect lexical collocations.

1.4 Purposes of the Study

The main purpose of the present study is to raise NEMs' collocation awareness via utilizing COCA. Considering the fact that the subjects in the present study were intermediate level students who had scored high in a previous CET Band-4 before, they could master grammar rules well but it is hard for them to choose right words to express themselves. The present study only focuses on lexical collocations. Therefore, the specific purposes of this study are to:

1. Identify the lexical collocation errors made by NEMs in KU in their writing;
2. Investigate the effects of utilizing COCA on raising learners' awareness concerning lexical collocations.

1.5 Research Questions

As mentioned before, the focus of the present study is on investigating participants' lexical collocation errors in their writing and using COCA for raising students' lexical collocations awareness. Therefore, the researcher intends to find out answers to the following two questions:

RQ 1. What are the typical lexical collocation errors found in the writing of non-English major EFL learners?

RQ 2. What are the effects of using COCA to raise students' lexical collocation awareness?

1.6 Significance of the Study

In China, though research reports and literature on the teaching and learning of vocabulary are very extensive, studies on collocations are still in need. After more than two years of intensive English training in college, many NEMs in KU have had a very good command of grammar and a large number of words. They can read and listen comparatively well but still have a difficulty in putting their ideas across to others. Many sentences made by them, in most cases, are grammatically correct but do not make sense or sound natural.

Becker (1995) provides a comment on those phenomena:

“We start with the information we wish to convey and the attitudes toward that information that we wish to express or evoke, and we haul out of our phrasal lexicon some patterns that can provide the major elements of this expression”.
(p. 72)

Becker’s remarks are very much to the point because he stresses the importance of collocation. Crowther (2001) puts the importance of collocation in this way: “perhaps, even more importantly, language that is collocationally rich is also more precise”. Actually, collocation is one of the factors responsible for Chinese EFL learners’ inadequate writing competence (Meng & Li, 2005).

Therefore, the present study attempts to find a way to help NEMs in KU to improve their English writing by effective use of English collocations. Through using COCA to search and collect collocations, and identify collocation errors, the researcher focuses on NEMs’ incorrect collocations in their production of language and tries to find efficient ways to raise their awareness of collocations. This would have some pedagogical implications. Collocation is a central aspect of vocabulary study, and it can provide a reliable guidance for language teaching and learning. The researcher hopes that the results from this study will shed some light on English language teaching and learning in Chinese colleges and universities so as to help the learners to achieve native-like competence and performance as much as possible.

1.7 Definitions of Key Terms

In this thesis, some specific terms should be viewed under the same light from the beginning to the end.

Among these specific terms, “English collocation” should be placed first, which has been defined by many researchers. Here, the researcher would take some into consideration and offer a working definition for “collocation”:

(1) Collocation

Collocation is viewed as the tendency of relationship of words’ mutual expectancy, one word to co-occur with one or more other words in a particular context, or several words which naturally or regularly go together through common usage. For example: *go to college, in the years to come, feel uncomfortable.*

(2) Lexical collocation

Lexical collocation refers to the predictable ways in which a noun, verb, adjective or adverb is combined with a word from another word class.

Besides “Collocation”, several other key terms are used in this study:

(3) COCA

The Corpus of Contemporary American English (COCA) is the largest and most recent freely-available corpus of English, and the only large and balanced corpus of American English. It was created by Mark Davies of Brigham Young University in 2008, and comprises more than 425 million of texts and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20

million words each year from 1990-2011 and the corpus is also updated once or twice a year (the most recent texts are from April 2011).

(4) Collocation errors

Collocation errors in this thesis refer to all the collocations which deviate from the norms of the target language. Collocations that cannot be found in COCA when set MI 3.0 with raw frequency at least 3 are considered collocation errors.

(5) Node

Each concordance line (word in a context) exemplifies a particular word or phrase that is being studied. This word or phrase is called the “node”. For example: When a learner queries the word “stimulate” in COCA, there are 3604 concordances showing how to use this node word “stimulate”. Such as: to *stimulate* economy, to *stimulate* interest.

(6) Collocate

A node is normally presented with other words to the left or the right and these words are called “collocate”. For example, for ‘high school’, ‘high court’, ‘high street’, and ‘so high’, ‘high’ is the node word and the other words, i.e. ‘school’, ‘court’ and ‘so’ are the collocates of the node word ‘high’.

1.8 The Overall Structure of the Study

The whole thesis is composed of five chapters.

Chapter One is the introduction, which states the research background. It provides a brief introduction to the study which focuses on the collocation

proficiency of the non-English majors in Kaili University. It also covers the research problems, purposes of the research, research questions, and significance of the study.

Chapter Two presents a literature review, which offers definitions and classifications of collocations, collocation competence and teaching of collocations. Also in this chapter, previous studies on English collocations and relevant theories on error analysis are presented.

Chapter Three provides the research methodology and its implementation, which is regarded as the research design. The design includes participants and data collection, and data analysis. Moreover, a summary of a pilot study is discussed to reveal merits and deficiencies of the methodology adapted.

In chapter Four, the results of the main study are presented and discussions on those results provided. Answers to Research Question 1 “What are the typical lexical collocation errors found in the writing of non-English major EFL learners at KU?” and Research Question 2 “What are the effects of using COCA to raise students’ lexical collocation awareness?” are offered.

In chapter Five, the research findings are summarized and the conclusion is drawn according to the results of the study. The pedagogical implications to collocation learning and teaching are presented. In addition, the limitations and suggestions for further research were described in details.

CHAPTER 2

LITERATURE REVIEW

In the past five decades, numerous researchers have devoted their time to the study of collocation in second language learners' written or spoken language. As stated in Chapter One, the main aim of the present study is to investigate participants' lexical collocation errors in their writing and using COCA to raise students' collocation awareness. To display what has been done in this field, relevant literature will be presented in this chapter. The discussion can be classified into four parts: 1) collocation; 2) error analysis; 3) previous studies of teaching English collocation in EFL contexts; and 4) teaching of collocation.

2.1 Collocation

2.1.1 Definitions of Collocation

The term “collocation” is defined by linguists and researchers in various ways and collocation is an interesting topic in contemporary linguistic research. However, there is still no widely agreed definition for collocation. Different scholars have slightly different ideas on what is meant by collocation so their definitions differ slightly from one another.

J.R. Firth, acknowledged as the father and pioneer of collocations, defines

collocation as “the company words keeps” or “the way words combine in predictable way” (as cited in Hill, 2000, p. 48). To Firth, collocational meaning is lexical meaning “at the syntagmatic level” (as cited in Gitsaki, 1999, p. 3). Collocation was later popularized by John Sinclair, another pioneer in the study of collocation. He provides a new way to look at the lexical structures. Sinclair’s conception stems from authentic language data. He defines collocation as, “the occurrence of two or more words within a short space of each other in a text” (Sinclair, 1991, p. 170). Besides, “collocation can be dramatic and interesting because unexpected, or they can be important in the lexical structure of the language because of being frequently repeated” (p. 170). There are three useful and practical terms in Sinclair’s description of a collocation: each concordance line exemplifies a particular word or phrase that is being studied. This word or phrase is called the “node”. A node is normally presented with other words to the left or the right and these are called collocates. The collocation can be counted and this measurement is called span. For example, we can find collocates within 4 words to the left and 4 words to the right for a given words.

Lewis (2000) defines collocation as “the way in which words co-occur in natural text in statistically significant ways” (p. 132), and “the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency” (Lewis, 2002, p. 8). Examples of collocations are: *submit a report*, *be aware of the risks*, and *on the other hand*. Similarly, Baker (1992) defines collocations as a tendency of certain words regularly co-occurred in a given language. Likewise,

Hill (2000) says that a collocation is a predictable word-combination. Nation (2001) points out that the term “collocation” is used to refer to “items which frequently occur together and have some degree of semantic unpredictability” (p. 317). Examples are *by the way, take a chance, and keep a diary.*

Some scholars take grammatical factors into account when defining collocation. Kjellmer’s (1987) definition of collocation is, “a sequence of words that occurs more than once in identical form in a corpus, and which is grammatically well structured” (p. 133). This definition indicates that collocations are lexically defined and grammatically restricted sequences of words. From Kjellmer’s stance, only sequences of two or more lexical words (some of them also incorporating function words) or sequences of one lexical word and one or more function words that recur in identical forms can be said to be collocations. Examples of collocation include: *in the years to come, brought about by, it is obvious that, must admit that, to apply to, and in the common experience of.* Cater (1998) states that collocation is a grouping of the words which naturally or regularly go together through common usage in language. His definition indicates that wrong collocations are clear markers outside of natural English. Cowie (1994) also takes grammatical structure into consideration. He suggests that “collocations are associations of two or more lexemes (or roots) recognized in and defined by their occurrence in a specific range of grammatical constructions”(p. 3169). To some extent, it seems that the randomly co-occurring words do not necessarily belong to a collocation unless they are grammatically well

structured.

Wray (2005) helps to distinguish “collocation” from “chunks”. He argues that a “chunk” is prefabricated, pre-determined, and readily observable while a “collocation” is by no means readily-made. It is not determined by logic or frequency, but by linguistic conventions (Lewis, 2002). In the meantime, there is a difference between the “English idiom” and “collocation”. An “idiom” is a fixed group of words which when taken together mean something different from the individual words when they stand alone. In other words, idioms have special meanings and the meanings must be learned as a whole, for example, *blow the gaff*; *the fourth estate*; and *black and blue*. Unlike “idioms”, the meaning of a collocation can usually be understood from individual words. In a word, as McCarthy (1990) states, collocation is a binding force between the words of a language which is distinct from fixed syntax of idioms and chunk. Nation (2001) states that when language users segment language for reception or production or to hold it in memory, they typically work with meaningful groups of items, called “chunks”, the size of chunks depends on the level of language proficiency the users have attained.

Furthermore, it is creative. Krishnamurthy (1997) suggests that the definition of regular or significant collocations is “Lexical items occurring within five words either way of the headword with a greater frequency than the law of averages would lead you to expect” (p. 70). This shows that what the linguist is interested in is the regular or significant collocation established on the basis of statistics by use of corpus

evidence. While the above definitions of collocations are given by authorities in language acquisition, Halliday and Hason (2001) stress the relation between vocabulary and discourse. They view collocation as one of the two aspects of lexical cohesion, "that is achieved through the association of lexical items that regularly co-occur" (p. 284). The co-occurrence of lexical items includes both syntagmatic relations and paradigmatic relations.

In addition to these definitions, Benson et al (1997) have a special idea on collocation. They suggest that collocation refers to not only the semantic habitual combinations, but also the grammatical ones, and define collocation in the introduction to the *BBI Combinatory Dictionary* as:

In English, as in other languages, there are many fixed, identifiable, non-idiomatic phrases and constructions. Such groups of words are called recurrent combinations, fixed combinations, or collocations. (1997: ix)

According to Benson et al (1997), a grammatical collocation is a type of construction that consists of a dominant word (verb, noun, adjective) and a prepositional or grammatical structure such as an infinitive or a clause. Examples of grammatical collocations are: *put up*, as in "*I can't put up with this any more.*" and *fill out*, as in "*You have to fill out your form.*" Lexical collocations normally do not contain prepositions, infinitives, or clauses. A lexical collocation is a type of construction where a verb, noun, adjective or adverb forms a predictable connection with another word. Benson et al (1997) distinguish collocations from idioms and free combinations. Meanings of collocations can be decoded from their constitute parts, which is in

contrast with idioms. The components of free combinations are substitutable. For instance, *condemn murder* is a free combination where the two elements are not bound specially to each other. They can occur with other words freely. Idioms, on the other hand, are frozen expressions whose meanings are opaque so we cannot figure out the meaning from the components such as *to scream blue murder* (to complain loudly).

As mentioned above, linguists can not agree on what a collocation really is. A number of factors are responsible for defining this term, among which are lexical co-occurrence, grammatical construction, co-occurrence frequency, and word choice. The grammatical construction and word choice can be identified by human intelligence while the lexical co-occurrence and frequency can be calculated by computers.

In trying to define what a collocation is, the present researcher is eclectic, combining the different definitions given by different linguists. “English collocation” is defined as the tendency of words’ mutual expectancy, or the co-occurrence of words in a particular context, or the cluster of several words which naturally or regularly go together through common uses, such as ‘*go to college*’, ‘*in the years to come*’, and ‘*feel uncomfortable*’. Considering that the subjects in the present study are intermediate students, they can master grammar well. The collocation proposed by the researcher in the thesis is limited to “lexical collocations”: the predictable ways in which a noun, verb, adjective or adverb is combined with a word from another word class. The researcher agrees with the point of view of Benson et al (1997), considering collocations to be not only the

semantic habitual combinations, but also the grammatical ones.

In defining collocations, frequently the contention is whether idioms are collocations. After considering the definitions of collocations offered by different linguists and researchers, the researcher regards idioms as consisting of words that occur in a rather fixed partnership and they are so tightly knit that they have virtually become independent lexical items (Benson et el, 1997), so idioms are distinct from collocations. Free combinations are not considered as collocations, because collocations are relatively fixed and with some degree of arbitrariness (e.g., Benson, 1989, Nesselhauf, 2003; Smadja & Mckeown, 1991), while the constituents in free combinations are not bound specially to each other (the elements of free combinations are substitutable). Taking the relationships between *strong / powerful* and *argument/tea/car* for an example (“√” means they can collocate):

	argument	tea	car
strong	√	√	
powerful	√		√

(as cited in Carter and McCarthy 1988, p. 34)

Figure 2.1 Arbitrariness of Collocation

As the figure suggests, people can say “*strong tea*”, “*powerful car*” but cannot say “*strong car*”, “*powerful tea*”. There is no logical explanation to the reason why “*strong*” cannot occur with “*car*” and “*powerful*” cannot occur with “*tea*”.

One of the important features of COCA is the possibility to measure Mutual

Information (MI) scores, according to Davies (2008) “typically, [MI] scores of about 3.0 or above shows [sic] a “semantic bonding” between the two words”. Church & Hanks (1990) note that pairs with scores above 3.0 can probably be considered collocations and below that free combinations. They give the detailed information of

MI:

“MI compares the probability of observing the joint probability of the two words x and y together with the probabilities of observing x and y independently (chance).

$$I(x, y) = \log_2 \frac{P(x, y)}{P(x) P(y)}$$

If there is a genuine association between x and y, then the joint probability P(x, y) will be much larger than chance P(x) P(y), and consequently I(x, y) > 0.” (Church & Patrick, 1990, p. 23).

The higher MI score the word combination has the stronger association between the two words. For instance: *soft drinks* (MI=6.67) is stronger than *soft voice* (MI=3.97). In other words, *soft* collocates with *drinks* is stronger than to collocate with *voice*.

McEnery et al (2006) affirm that “collocations with high MI scores tend to include low-frequency words” (p. 57), and infrequent collocations do not appear particularly interesting for language teaching. Since the English language has thousands upon thousands of words, frequent collocations should be taught or students would be learning language items which they would probably never use. Another criterion established to decide whether a word combination is a collocation is to check frequency. Clear (1993) adopts the threshold value of three,

discarding pairs appearing fewer than three times in corpus. Likewise, in the present study when a word combination found in COCA with a raw frequency lower than three, it is not considered a collocation. Thus, only word combinations which have MI scores higher than 3.0 and raw frequency higher than 3 in COCA are classified as collocations. For the sake of convenience, word combinations which cannot be found in COCA when set MI 3.0 with raw frequency at least 3 are not collocations and would be filtered out.

From the researcher's stance, "chunk" is the broadest term. "Chunks" are meaningful units. Collocations, idioms, and free combinations are all "Chunks". Language can be accounted for by the storage of chunks in long-term memory (Nation, 2001). According to learners' proficiency level, a chunk can be letters, words, morphemes, collocations, or even a song, a poem. Different learners view "chunks" in different ways, depending on the way they "chunking" words--- the way they look at words as a meaningful unit to store in their long term memory. A good example is the memory of a series of numbers such as 2471512. One learner may look at it in this way: 24 (a day), 7 (a week), 15 (half-month), 12 (a year), but another learner may think 247 (room number), 1512 (cell phone number).

2.1.2 Collocation Competence

Knowledge of collocation is part of native speakers' competence. Native speakers can fluently express themselves and speak at a relatively fast speed because they own enough ready-made language chunks in their mental lexicons.

Their reading and listening comprehension is ‘quick’ too because they can constantly recognize ‘chunks’ of language. Due to having stored a large amount of collocation items, native speakers can communicate and think more efficiently, and the way of processing also contributes to the main difference between native speakers and non-native speakers. Kjellmer (1991) argues that:

“In building his utterances, he [the native speaker] makes use of large prefabricated sections. The learner, on the other hand, having automated few collocations, continually has to create structures that he can only hope will be acceptable to native speakers. His building material is individual bricks rather than prefabricated sections.” (p. 124)

In the English language, combinations of words occur more often than expected by chance. Why do we say we go somewhere ‘by car’ or ‘by train’ but ‘on foot’? Why you *make a cup of tea*, but *do your homework*? The reason is ‘collocation’. Knowing the ‘meaning’ of a word not only requires knowing its dictionary definition, we must also know the type of words with which it is often associated. Hill (2000) explains that most learners with ‘good vocabulary’ have problems with fluency because their ‘collocational competence’ is very limited, which is similar for Chinese university English learners (Meng & Li, 2005). As a result, it is common that even at higher levels of proficiency, students lack conciseness and precision of expression. Lack of collocational competence leads to errors. Students “tend to create longer utterances because they do not know collocations which express precisely what they want to say” (Hill, 2000, p. 49).

Lewis (2000), in his ‘lexical approach’ theory, states that fluency in a foreign

language is conditioned by the acquisition of a number of pre-fabricated chunks. The richer in collocations the learner's lexicon is, the higher precision, accuracy, coherence and authenticity of his/her speech.

Oxford Collocations Dictionary (2002) states that:

“For the student, choosing the right collocation will make his speech and writing sound much more natural, more native speaker-like, even when basic intelligibility does not seem to be at issue... A student who chooses the best collocation will express himself much more clearly and be able to convey not just a general, but something quite precise” (p. vii).

In conclusion, the main difference between native speakers and non-native speakers is that native speakers own a larger number of collocations to draw on for their spoken and written communication needs. Collocation competence refers to the ability to choose and use word combinations and expressions to make one's speech and writing more natural. Collocation competence includes the ability to recognize collocation during reading and listening, the ability to use collocation in spoken and written communication, and the ability to identify collocation patterns. To know how to classify collocation is a part of collocation competence, too.

2.1.3 Classification of Collocations

Just as different linguistic experts have defined collocation from different points of view, the classification of collocations has been conducted from various perspectives. Proposals have been made by previous linguists to categorize collocations according to different viewpoints: collocation as a continuum; collocation restriction; collocation strength; word frequency of collocation; and

characteristics and function of collocation.

2.1.3.1 Collocation as a continuum

As mentioned in the previous section, according to Howarth's (1998), collocations can be considered as a continuum. In his categorization model of lexical collocations, which is based on how closely the components in a collocation hold together, the collocational continuum contains four categories: free combinations, restricted collocations, figurative idioms, and pure idioms.

Kjellmer (1990) argues that English words are spread over a continuum with items whose contextual company is entirely predictable at one end, and items whose contextual company is entirely unpredictable at the other. Lewis (2000) states, there are very few 'strong' collocations because most collocations are found in the middle of the continuum.

In the present study, collocations are viewed as a continuum with idioms at one end and freely occurring words at the other.

weak	medium-strength	strong	fixed
free combination	lexical collocation	grammatical collocation	idioms

(as cited in Malligamas & Pongpairroj, 2005)

Figure 2.2 Collocation continuum

2.1.3.2 Collocation restriction

Carter (1998) divides collocations into four categories, according to the degree of collocation restrictions:

1) The 'unrestricted', in which collocates freely occur with a number of lexical items, and the elements are substitutable, e.g. *have a headache/a haircut/ breakfast/ a bath; take a look/a holiday/a rest/a letter/time/notice/a walk.*

2) The 'semi-restricted', in which the number of adequate substitutes which can replace the elements of collocation is more limited, e.g. *without doubt/uncertainty/suspicion.*

3) 'Familiar' collocations, whose elements collocate on a regular basis, e.g. *unrequited love, lukewarm reception.*

4) 'Restricted' collocations, which are fixed and inflexible, and tend not to allow for substitution for elements, e.g. *dead drunk, pretty sure.*

2.1.3.3 Collocation strength

Hill (2000) distinguishes collocation into four categories:

1) Unique collocations: *foot the bill, shrug one's shoulders.* They are unique collocations because *foot* is used as a verb and *shrug* is not used with any other nouns.

2) Strong collocations: *rancid butter, trenchant criticism.* There are very few other things that can be *rancid* or *trenchant*.

3) Medium-strength collocations: *hold a conversation, make a mistake, a major operation.*

4) Weak collocations: *tall man, loud noise, red car, big house, nice day,*

good chance.

Based on Hill's (2000) collocation categories, it can be concluded that strong collocations are partnerships which are tightly linked. They may be frequent or comparatively rare. Weak collocations occur between two common words, each of which may co-occur with many other words.

2.1.3.4 Word frequency of collocations

Sinclair (1991) divides collocations into two categories: the 'upward' and 'downward' collocations.

1) The "upward" collocations consist of words which habitually collocate with the words more frequently used in English than they are themselves. For example, *look* collocates with *after, for, on, out, to, up*, all of which are more frequent words than *look*. The word *back* collocates with *at, down, from, into, on*, all of which are more frequent words than *back*.

2) The 'downward' collocations are words which habitually collocate with words that are less frequent than they are. For example, words *arrive, bring* are less frequently occurring collocates than *back* and words *alive* and *sharp* are less frequently occurring collocates than *look*.

Sinclair (1991) makes a sharp distinction between those two categories, claiming that the elements of the 'upward' collocation (mostly prepositions, adverbs, conjunctions, pronouns) tend to form grammatical frames while the elements of the 'downward' collocation (mostly nouns and verbs) by contrast give a semantic analysis

of a word.

2.1.3.5 Characteristics and functions of collocations

Benson et al's (1997) model of 'grammatical and lexical' collocations is theoretically sound and pedagogically useful. Grammatical collocations usually consist of a noun, an adjective or a verb plus a preposition or a grammatical structure such as 'to+infinitive' or 'that-clause', e.g. *by accident, to be afraid that*. Lexical collocations do not contain grammatical elements, but are combinations of nouns, adjectives, verbs, adverbs (Bahns, 1993).

Benson et al (1997) define collocation as specified, identifiable, non-idiomatic, recurrent combinations. In their dictionary they divide them into two groups: grammatical and lexical collocations. The first category consists of the main word (a noun, an adjective, or a verb) plus a preposition or 'to+infinitive' or 'that-clause' and is characterized by 8 basic types of collocations. In Table 2.1, these types are presented:

Table 2.1 Grammatical Collocation Types

Types	Examples
G1 = noun + preposition	<i>put forward;</i> <i>blockade against;</i>
G2 = noun + to-infinitive	<i>He was a fool to do it;</i> <i>They felt a need to do it;</i>
G3 = noun + that-clause	<i>She took an oath that she would perform the duty;</i> <i>We reached an agreement that she would represent us in court;</i>
G4 = preposition + noun	<i>by accident;</i> <i>in agony;</i>
G5 = adjective + preposition	<i>fond of children;</i> <i>hungry for news;</i>
G6 = adjective + to-infinitive	<i>it was necessary to work;</i> <i>it's nice to be here;</i>
G7 = adjective + that-clause	<i>She was afraid that she would fail the examination;</i> <i>It was imperative that I be here;</i>
G8 = different verb patterns in English (mainly refers to 19 kinds of English verb drills) <i>Verb+ing; verb + to-infinitive; verb + bare infinitive; and other. Etc.</i>	<i>They enjoy watching TV;</i> <i>They began to speak;</i> <i>We must work;</i>

Note: Adapted from Benson et al. (1997) pp. xvi – xxviii

Lexical collocations do not contain prepositions, infinitives or relative clauses but consist of nouns, adjectives, verbs and adverbs. Originally, Benson et al (1997) classified lexical collocation into seven categories, since L1 Verb (which means creation/action) + noun category and L2 Verb (which means eradication/cancellation) + noun category proposed by Benson, et al (1997) are similar, the researcher combines these two categories into one L1 (V+N) category. There are 6 types of them which are presented in Table 2.2:

Table 2.2 Lexical Collocation Types

Types	Examples
L1 Verb+noun	<i>to cancel an appointment;</i> <i>to reject an appeal;</i> <i>to reach a verdict;</i>
L2 Adjective +noun	<i>strong tea;</i> <i>reckless abandon;</i> <i>a crushing defeat;</i>
L3 Noun+verb (naming the activity which is performed by a designate of this noun)	<i>bombs explode;</i> <i>bees sting;</i> <i>alarms go off;</i>
L4 quantifier + noun (group or units of thing)	<i>a swarm of bees;</i> <i>a pack of dogs;</i> <i>a piece of advice;</i>
L5 Adverb+adjective	<i>sound asleep;</i> <i>closely acquainted;</i> <i>hopelessly addicted;</i>
L6 Verb+adverb	<i>run rapidly;</i> <i>apologize humbly;</i> <i>argue heatedly;</i>

Note: Adapted from Benson et al. (1997) pp. xxx – xxxv

The collocation classification system to be used in the present study is from Benson et al (1997). The reasons for this preference are:

- 1) The classification system is clear.

As listed in Table 2.1 and 2.2, Benson et al (1997) classify English collocations into two major groups: lexical collocation and grammatical collocations. Lexical collocations do not contain prepositions, infinitives or relative clauses but consist of various combinations of nouns, adjectives, verbs and adverbs. There are seven types.

Grammatical collocations consist of the main word with a preposition or “to infinitive” or “that-clause” and are characterized by eight basic types of collocations and 19 kinds of English verb drills. In order to make the demarcation between each subcategory clear, each of its categories is presented in details and with examples.

2) The classification system is comprehensive.

Benson et al (1997) provide a very comprehensive list of lexical and grammatical collocations. It includes 7 lexical and 26 grammatical collocation types, which have been adopted by many researchers to investigate L2 collocations (Li, 2008; Fan, 2009; Kuo, 2009; Hsu, 2010). In most cases, the investigations into L2 collocations are narrowed down to collocations of particular structures. For examples, Bahns and Eldaw (1993), Al-Zahrani (1998), and Howarth (1998) examine verb + noun collocations, Granger (1998) and Lorenz (1999) study the adverb + adjective collocations. Some researchers adopt these types to design tests to examine learners collocation competency (Abadi, 2003; Keshavarz and Salimi, 2007; Al-Sibai, 2009).

3) The classification system can be taught with explicit collocation teaching.

Learners can be trained to use Benson et al' s classification as a model to recognize collocations during reading and listening, to identify collocation patterns, and to choose which collocation can be used in spoken and written communication.

As for categorizing collocations, Wei (1999), Bonk (2000), and Zughoul & Hussein (2003) have recognized Benson et al' s model of grammatical and lexical collocations as theoretically sound and pedagogically useful. Wei (1999) states that

Benson et al 's (1997) classification systematically incorporates syntax into a semantic and lexical construct, thus encompassing a wide range of data. The broader scope is consistent with the pervasive nature of collocation and the inclusion of syntax makes the model more amenable to teaching.

According to the characteristics and functions, collocations are divided into two main categories: grammatical and lexical collocations. The current study only focuses on lexical collocations. The Lexical Collocation Types in Table 2.2 are useful for understanding EFL students' linguistic performance and helping them improve it. It will not only be used as a model to provide guidance in explicit teaching of collocation awareness and skills to identify collocation patterns, but also be used to count and calculate lexical collocations errors.

2.2 Error Analysis

Error analysis in SLA was established in the 1960s by Stephen Pit Corder and colleagues. It is a type of linguistic analysis that focuses on the errors learners make. There is a general agreement over the division of interlingual errors and intralingual errors by linguists (James, 1998). Errors that results from mother-tongue influence is called interlingual errors, which are caused by language transfer. Errors that results from faulty or partial learning of the target language, which refers to those produced in using target language in own terms is called intralingual errors. Speaking of error, it is produced in foreign / second language learners' spoken or written production. If we do

not pay more attention, it will lead to the Chinglish and unidiomaticity in L2 production which is caused by incompetence in using collocations. Experts distinguish between errors and mistakes. Error, which can be defined as “a deviation from the norms of the target language” (Ellis, 1994, p. 51), can be considered as a lack of English knowledge, showing a lack of linguistic competence. Those errors are likely to occur repeatedly and not recognized by the learner. Mistakes occur when a learner fails to perform his/her competence due to slips of tongue, memory limitations or lack of linguistic planning and strategy. James (1998) points out that an error is not a self-correctible and unintentionally deviant phenomenon of language while a mistake is a self-correctible and either intentionally deviant or unintentionally deviant phenomenon.

It is very clear that errors and mistakes belong to different kinds of failure of language. In the current study, the researcher adopts Ellis' (1994) definition: an error is “a deviation from the norms of the target language” and will not make the distinction between mistakes and errors. Errors in the thesis refer to all the deviations from the norms of the target language.

Error Analysis is “the process of determining the incidence, nature, causes and consequence of unsuccessful language performance” (James, 1998, p. 1). Error Analysis focuses on the errors made by the learners, and particularly researchers are supposed to classify errors according to their causes and sources.

Corder (1981) points out that in the process of error analysis, concentration only on superficially ill-formed sentences (overtly idiosyncratic) is not enough. Those that are

well-formed but inappropriate in the context must also be dealt with. Such sentences are called covertly idiosyncratic by Corder and cannot be interpreted normally in contexts. Both overtly idiosyncratic and covertly idiosyncratic sentences have to be analyzed. Corder proposes the procedure for error analysis which includes three stages:

1. Data collection: Recognition of idiosyncrasy;
2. Description: Accounting for idiosyncratic dialect;
3. Explanation (the ultimate object of error analysis).

Based on Corder's (1981) procedures of error analysis, Chow (2006) employs five steps to deal with grammatical errors in her study: (1) identify errors; (2) categorize errors; (3) analyze errors; (4) evaluate errors; and (5) correct errors. The present study will take the following steps in using COCA to analyze learners' collocation errors:

- Step 1. Collection of a sample of learner's language and identification of errors

This stage is a process of collection of a sample of learner's language and then identifying errors. The sample of the learner's language can be collected cross-sectionally or longitudinally. In addition, natural samples are generally preferred compared with elicited data. The samples of my study are collected from Chinese learners' writings. It should be carried out according to standard written dialects (concordances in COCA). Corder (1981) offers two ways to identify errors: authoritative interpretation and plausible interpretation respectively. The former detects errors by interviewing the learners. If the learners are not available for the consultation, the latter can be applied to give an interpretation on the basis of the

forms and the linguistic and situational contexts.

Step 2. Comparison and classification of errors

As soon as errors are identified, the next step is to describe them properly. The description of learners' errors involves two aspects: comparison and classification. First, the researcher can compare the learners' errors with the reconstruction of the sentence in the target language (COCA), and then, errors can be classified into different types in accordance with Benson et al's (1997) lexical collocation types.

Step 3. Explanation and evaluation of errors

This stage is concerned with identifying the sources of errors and evaluating errors, which is very important because some errors can reflect learners' attempts to perform the task. Errors are fundamentally involved in the learning process at certain stages and making errors is unavoidable during L2 learning. "Evaluating an error means assigning relative value and standards to it, that is, deciding which error merits attention" (Yi, 2004, p. 86).

Through identifying collocation errors in NEMs' writing and collocation tests in KU, the researcher focuses on finding efficient ways to raise NEMs' collocation awareness in KU. This would have some pedagogical implications. It would help English teachers in KU to know students' weaknesses in English learning and would provide more reliable guidance for instruction in English language teaching and learning. The researcher hopes that the results from this study will shed some light on how to improve English language teaching quality and enhance learning efficiency

and to help NEMs in KU to achieve native-like competence and performance.

2.3 Sources of collocation errors

There has been a great concern among researchers about the reasons why EFL students frequently make collocational errors (Bahns,1993; Farhal & Obiedat,1995; Al-Zahrani,1998; Liu,1999; and Meng & Li, 2005). Studies have shown that the common sources of collocation errors are related to the L1 interferences, overgeneralization, and shortage of collocational knowledge.

In terms of the L1 interference, learners' first language influences their production of collocations. It has been found in Chinese EFL learners' writings that learners would use *eat tea* instead of *drink tea*, *quick speed* instead of *fast speed*, and *mother school* instead of *alma mater*. They make these errors because they translate Chinese into English according to the Chinese linguistic convention. These expressions are understandable in Chinese but are not acceptable in English. Bahns (1993) investigated Polish and German EFL learner's performance in English collocations. Based on this study, Bahns concluded that the majority of collocational errors can be traced to L1 influence. A number of subjects provided "*win money*", "*make attention at*", "*drive a bookshop*", and "*finish a conflict*" for the target collocation "*make money*", "*run a bookshop*", "*pay attention to*", and "*solve a conflict*". Similar conclusion was made by Bahns and Eldaw (1993), Farhal and Obiedat (1995), Al-Zahrani (1998), Liu (1999), Meng and Li (2005), and Fan (2009). They proclaimed that many EFL learners' collocational errors are caused by their L1

interference.

With respect to overgeneralization, Liu (1999) points out that overgeneralization is the creation of a deviant structure in place of two regular structures on the basis of students' experience of the target language. For example, students who learnt collocations: *a missing boy*, *a knowing smile*, probably consider that "verb + ing" functions as an adjective, so such collocation errors appeared in their writing: *attracting feature*, *amusing look*, *satisfying smile*. Similar conclusion was reported by Meng and Li (2005).

With regard to the shortage of learners' collocational competence due to the lack of collocational knowledge, EFL learners may think that words such as *make*, *do*, and *take* are de-lexicalized verbs so they can replace one another freely (Liu,1999). Therefore, the participants made errors such as *make some measures* instead of *take measures*, *do plans* instead of *make plans*. They may ignore rules and restrictions and use incorrect vocabulary items or structures. Use of a coined word or use of a synonym for a lexical item in a collocation is seen as a "straightforward application of the open choice principle" (Farghal & Obiedat, 1995).

Howarth (1998) compared written corpora of EFL learners with native English speakers, and found that EFL learners made a lower percentage of conventional collocations but a higher percentage of deviant combinations. e.g. *perform a project*, *write a table*, *pay effort*, and *reach finding*. He reported that both groups have insufficient knowledge of collocations in general. Other researchers such as Bahns

and Eldaw (1993), Farghal and Obiedat (1995), and Fan (2009) also found that L2 learners had a big gap between their receptive and productive knowledge of collocations.

To sum up, EFL students make collocational errors in their writing because of the interference of their L1, overgeneralization of L2 rules, and their shortage of collocational knowledge. These can be used to explain why NEMs in KU make unacceptable collocations in their writing and collocation tests. Analysis of the sources of collocation errors can help NEMs in KU better understand the target language and thus they can avoid making collocation errors and produce natural English.

2.4 Previous Studies on Teaching English Collocations in EFL Contexts

In this section, several empirical studies on teaching English collocation in EFL contexts are reviewed. Bahns and Eldaw (1993) conducted a study of the advanced German EFL students' productive knowledge of English collocations consisting of a verb and a noun by using a translation and a cloze task. Fifty-eight German EFL college students who participated in the study were divided into two groups. Twenty-four students in the first group were required to take a cloze test with 10 sentences, each of which had a verb-noun collocation with the verb missing. The other twenty-four students in the second group were required to take a

German-English translation test, which consisted of 15 sentences. In the end, it was found that in both tests only half of the students' responses were acceptable English collocations. Students were given more freedom to paraphrase in a translation test. As a result, they produced more than twice as many errors in the translation of verbal collocation as in the translation of general lexical words. Therefore, they concluded that for advanced German students, collocations present a major problem in the production of correct English. Learners are not aware of collocations as a potential problem in language learning. EFL teaching should concentrate on those collocations which cannot readily be paraphrased.

Farghal and Obiedat (1995) examined Arabic EFL learners' command of collocational knowledge by adopting two measures: an English fill-in-the-blank test and an Arabic translation test involving twenty commonly used collocations related to the topics such as weather, color, and food. They had thirty-four intermediate learners taking the fill-in-the-blank test and twenty-three advanced learners taking a translation test. The correct collocations obtained were 18.3% and 5.5% respectively. In the end, Farghal and Obiedat analyzed these mis-collocations and declared that EFL teaching should focus on collocation especially, which are unpredictable from learners' first language.

Similarly, Gitsaki (1999) measured two hundred seventy-five Greek junior students' English collocational knowledge by using an essay task, a blank-filling task, and a translation test. Students were divided into three groups according to their language proficiency levels: post-beginning, intermediate, and post-intermediate.

Gitsaki noted that intermediate students' collocational knowledge does not increase alongside the grammatical competence. And, grammatical collocations seem to be easier for students to acquire than lexical collocations.

Wei (1999) reported that one important undervalued aspect of productive vocabulary is collocation—the way in which the words are combined with one another. To move from receptive to productive vocabulary, teachers need to teach collocations and students need to learn a wide variety of ways that words collocate with each other. Raising students' awareness of collocations will motivate them to find collocations from reading, dictionaries and communication with native speakers. All these will enhance their language development. Some effective activities and techniques for teaching collocations have been presented in his study. Such as peer correction, sentence making, and RDRR activity. RDRR refers to the four steps: Read: study collocations and examples you select in books; Don't read: use collocations to make sentences without looking at collocation examples; Reread: reread the examples again to make necessary revision or correction; and Reuse: review and reuse the collocations regularly.

Lien (2003) explored the effects of collocation instruction on reading comprehension. She designed and carried out a quasi-experimental study at a national university in central Taiwan over a four-week period. Subjects were eighty-five Taiwanese college English majors from three academic levels (i.e., sophomores, juniors, and seniors). Before reading three different articles of a similar

length and difficulty, the students of three academic levels received three types of treatment —vocabulary instruction, collocation instruction, and no instruction—and took three immediate reading comprehension tests consisting of ten short essay questions each. Lien's findings suggested that EFL learners' reading comprehension somewhat increased if receiving pre-reading word instruction and their performances on reading comprehension improved along with their academic levels. However, there was no significant difference observed in the participants' readings comprehension performances among the three academic levels as a result of different treatments even though the sophomores (i.e., the lowest-level subjects) were found to react best after collocation instruction.

Mallikamas and Pongpairoj (2005) conducted a study to investigate Thai learners' receptive and productive knowledge of English collocations and analyzed their problems in the use of three types of collocations: lexical, grammatical and bound. One hundred and fourteen Thai freshmen in the Faculty of Arts at a Thai University took part in the study. Two sets of tasks were developed to test the subjects' receptive and productive knowledge of the collocational relationships between words. Each set was composed of three parts. In each part, the subjects were tested for knowledge of three types of collocations: lexical, grammatical and bound. Quantitative analysis was used to identify whether any type of collocations posed more difficulty than the others. The results indicated wide-ranging problems in Thai learners' collocational knowledge. While grammatical collocations posed a problem

for learners in multiple choice and gap-filling tasks, lexical and bound collocations caused more problems in reception than in production. The researchers strongly advocated the lexical approach to help develop Thai learners' collocational knowledge.

While the importance of collocation acquisition has been recognized for a long time, in Japan, vocabulary learning has often been limited to memorizing lists of words in isolation. Zorana (2008) found some explicit teaching of collocational activities that are effective with Japanese college students and presented them in his study. Such as: find intruder, correct a mistake, focus on paraphrases, collocation bingo, and collocation translation. Webb and Kagimoto (2010) investigated the effects of three factors (the number of collocates per node word, the position of the node word, and synonymy) on learning collocations. Forty-one Japanese college students in five second-year EFL classes at two colleges were selected according to their scores on a pre-test, which was administered one week before the experiment, measuring productive knowledge of target collocations. The pre-test used a productive translation format in which the L1 meanings of the collocations cued the sixty L2 collocations. Only participants who were unable to demonstrate any knowledge of the target collocations took part in the treatment. A posttest using the same format as the pre-test and the same target collocations was administered to participants. Five 12-item productive translation tests were presented in a different order from the pre-test. Each of the five 12-item tests was completed immediately after each set of

collocations was learned in glossed sentences. The glossed sentences were based on sentences in the British National Corpus and the Bank of English. The collocations and their L1 meanings were written in bold above each sentence, and the target collocations were also shown in bold in the sentences. The participants were given 3 minutes to learn each set of collocations. The results showed that more collocations were learned as the number of collocates per node word increased, the position of the node word did not affect learning, and synonymy had a negative effect on learning.

In order to know whether words could have been taught in collocations, Hodne (2009) conducted a study to investigate word combinations in two textbooks which were used in upper secondary schools in Norway through using COCA. The results turned out that the integration of collocations in vocabulary teaching is feasible and it is beneficial for learners. This study provided a new insight into the treatment of collocations in language material. It suggested the types of collocations that can be taught, and showed how these can be integrated in vocabulary exercises and teaching.

Hsu (2010) carried out a study to investigate the effects of direct collocation instruction on Taiwanese college English majors' reading comprehension and vocabulary learning. Three groups of Taiwanese college English majors were divided according to their academic levels. Each group received different collocation instruction (i.e. single-item vocabulary instruction, lexical collocation instruction, and no instruction), vocabulary tests, and reading comprehension tests. Through nine weeks'

collocation instruction, the result turned out that direct lexical collocation instruction and using first language glosses as a vocabulary-facilitating activity improved the subjects' vocabulary learning, promoted the subjects' performance on the three recall tests and Taiwanese colleges made significant progress in their reading comprehension as long as they received collocation instruction. This study suggested that direct collocation instruction could be a worthy option for exploration in teaching.

It is obvious that some researchers just studied college students (Bahns and Eldaw, 1993; Mallikamas and Pongpairoj, 2005; Webb and Kagimoto, 2010; Hsu, 2010) while others investigated middle school students (Gitsaki, 1999; Hodne, 2009). Previous studies have all shown that EFL learners are far incompetent in the use of English collocations. Research findings support that it is necessary to teach collocations in classrooms and that teaching collocations has positive effects on the growth of learner's collocational competence (Wei, 1999; Hodne, 2009; Webb and Kagimoto, 2010; Hsu, 2010). In order to improve learners' English language proficiency, collocation teaching should be stressed. In China, although research reports and literatures on the teaching and learning of vocabulary are very extensive, little research has been done on collocation errors made by learners who are NEMs and there are few empirical studies investigating the learners' difficulties in collocations or raising learners' collocation awareness to improve students' collocational competency.

English collocations have always been a main obstacle for the Chinese English

learners to achieve native-like proficiency (Liu, 1999; Chen, 2002; Meng, & Li, 2005; Fan, 2009). All these give inspirations to the researcher to conduct a study to explore the collocation errors in Chinese English learners' writings and raising students' collocation awareness via utilizing corpora. Thus, this study investigated NEMs' collocation errors in their writing in KU, explored their difficulties in the use of English collocations, and tried to find efficient ways to raise students' awareness of collocation. Hopefully, this would contribute to a better understanding of English learners' difficulties in collocations and shed more light on English language teaching and learning in Chinese colleges.

2.5 Teaching of Collocations

Having a large store of vocabulary is not enough for learners to produce natural English. A student may supposedly have a store of 'large size of vocabulary' but still produces 'unnatural' English due to lack of collocational knowledge. Nation (1990) commented that productive knowledge of vocabulary extends beyond the receptive knowledge to pronunciation, spelling, structures and collocation. Wei (1999) stated that to move from receptive to productive vocabulary, teachers need to teach collocations and students need to learn a wide variety of ways that words collocate with each other. The process of acquiring collocation is a laborious process through 'years of study, reading, and observation of the language' (Mackin, 1978, p. 151).

How could foreign language learners make their spoken and written

communication sound more natural? How could they achieve native-like English proficiency? In order to help learners produce natural English, collocations should be included when teaching all English skills. Collocations are too numerous to be taught in the classroom. While teaching collocation, teachers need to bear in mind that not all of the collocations have to be learned. Emphasis needs to be placed on building students' consciousness of how words work in combination with one another, so they can continue developing their collocational competence after class. The categories of collocation (Benson, 1997 (Table 2.2)) are useful for understanding EFL students' linguistic performance and helping them improve it. Clear and explicit presentation need to be used to provide guidance in building students' awareness. Lewis (2002) stated that the idea of collocation is a very powerful one in helping learners maximize the value of the language to which they are exposed, but they need help in identifying the powerful and useful partnerships in a text.

Nation (2001) believed that with the help of vocabulary learning strategies, a large amount of vocabulary could be acquired and these strategies are useful for students of different language levels. He proposed a taxonomy of various vocabulary learning strategies which is divided into three general classes of 'planning', 'source' and 'processes', each of which is divided into a subset of key strategies.

Table 2.3 Types of Vocabulary Learning Strategies (Nation, 2001, p. 218)

General class of strategies	Types of strategies
Planning: choosing what to focus on and when to focus on it	Choosing words Choosing the aspects of word knowledge Choosing strategies Planning repetition
Sources: finding information about words	Analyzing the word Using context Consulting a reference source in L1 or L2 Using parallels in L1 and L2
Processes: establishing knowledge	Noticing Retrieving Generating

Based on Nation's (2001) vocabulary learning strategies, for raising NEMs' awareness of lexical collocation in KU, under "planning", teachers and students chose content words to learn, chose lexical collocations to focus on, and planned repetition. Under "sources", students learnt collocation in contexts to notice how words co-occur together and what words to go with a given word, found information about collocations via utilizing COCA. For "processes", learners established collocation knowledge via receiving more input in classes, recalling items previously met, self or pair correction by using COCA to enhance accuracy, and paid attention to both receptive and productive knowledge of collocation.

As to the way in which new words are acquired, Hatch & Brown (1995) suggest five steps: (1) encountering new words; (2) getting the word forms; (3) getting the word meanings; (4) consolidating word forms and meanings in memory; and (5) using

the words. They provide examples of how the steps can be achieved. Learners will encounter new words, for example, on the Internet, television, books, movies, etc. Getting the word form is the step when students connect the sound with the word form, which can be done by reading the word aloud. As a way to get the word meaning, students can use bilingual or monolingual dictionaries or guess from the context. Consolidating word forms and meanings in memory can be achieved by doing a number of exercises such as matching or translating words, filling in gaps, and so on. Finally, to learn the uses of a word, students might be asked to write sentences or texts or consult concordances to check how the word is used. For acquiring collocations, as we learn new words by the means of collocation, these five-steps can be applied.

The instruction of learning strategies can help EFL learners become more efficient in their learning. In addition, “skill in using learning strategies assists students in becoming independent, confident learners” (Chamot, 1999, p. 1). Besides learning strategies, approaches such as techniques, exercises, and activities can be used to teach collocation (Wei, 1999):

1. Techniques : The teacher should focus learners’ attention on the relationship between words that often appear together, and help them to identify different types of collocations.

2. Exercises : Exercises should be designed to reinforce learning. Boonyasaquan (2006) suggested giving a variety of collocation exercises to the

students to challenge their ability.

3. Activities : Activities need to be interactive and student-centered to encourage students' interactive and maximize feedback. Examples of activities include underlining collocations in a text, concordances, incomplete phrases, finding pairs of collocations arranged randomly, and matching. These activities can be used to raise students' collocation awareness. A typical collocation activity gets learners to match collocates with given items. Paul Nation and Jonathan Newton suggest that collocation activities can be done as group or class activities with learners drawing on their differing experience to suggest collocates for a given word (as cited in James & Thomas, 1997). Teaching and learning the most frequent collocates of a node and its collocation-related range of meanings are the two activities strongly recommended by Nation (2001). By selecting items relevant to the topic of the students' assignment and focusing their attention to common word combinations, teachers can help make students' writing more accurate and natural.

It can be concluded that there are many interesting methods of teaching collocations. It is teachers' responsibility to decide what methods are appropriate for their students. The major objective is to raise the learners' awareness of the importance of collocation. In the present study, the techniques, exercises, and activities mentioned were used to raise students' awareness of collocations.

CHAPTER 3

RESEARCH METHODOLOGY

The preceding chapter presented the theoretical underpinning for the current research study. This chapter provides an overview of the research methodology of the study. It discusses the research design, the population, the participants, the instruments, the data collection, the data analysis, and pilot study. The researcher is interested in answering the following questions:

RQ 1. What are the typical collocation errors found in the writing of non-English majors in KU?

RQ 2. What are the effects of using COCA to raise students' lexical collocation awareness?

3.1 Research Design

3.1.1 Phase I Error analysis

Two writing tasks were given to one hundred and fifty intermediate NEMs in KU. After the participants had completed the writing tasks, 300 essays (150 participants \times 2 essays) were collected as a learner corpus. All lexical collocations to be found in the learner corpus as defined in Chapters 1 and 2 were underlined and compared with COCA for detecting and analyzing collocation errors. Twenty

participants who made the most frequent collocation errors were participants for a semi-structured interview, which designed to collect information about the reasons students made errors. The results of error analysis and the interview were used to discuss and answer RQ1.

3.1.2 Phase II A quasi-experimental study

To arrive at the answer to RQ 2, a quasi-experimental study in nature was conducted. Thirty-nine NEMs in KU who failed in CET Band-6 in December, 2010 took part in the instructional treatment. A pre- and post test was employed. Quasi-experimental research is a type of experimental research. Thomas (2003) argues that the most important characteristic of a quasi-experimental research study is to deal with the phenomenon of cause and effect.

To increase the validity of the study, theoretical and methodological triangulations were used. In this study, theoretical triangulation involved error analysis, semi-structured interview, and collocation instruction design. Methodological triangulation involved using quantitative and qualitative methods to collect data. The advantage of employing methodological triangulation mainly lies in that it can be used to address different but complementary questions within a study and enhance the interpretability for the research outcomes (Robson, 2002). In this study quantitative and qualitative methods were implemented to examine the effects of collocation instruction. Data collection was conducted within an 8-week instruction period. The class lasted fifty minutes every week. In order to collect the data on the effects of using COCA to raise students' lexical

collocation awareness, the pre- and post test collocation scores, a semi-structured interview, and a questionnaire were used. Data from the questionnaire and semi-structured interview were examined to investigate the students' attitudes towards using COCA to learn lexical collocation. And data from students' pre- and post test were collected and used to investigate improvement on students' collocation performance.

Table 3.1 Summary of Instruments or Means to Answer Research Questions

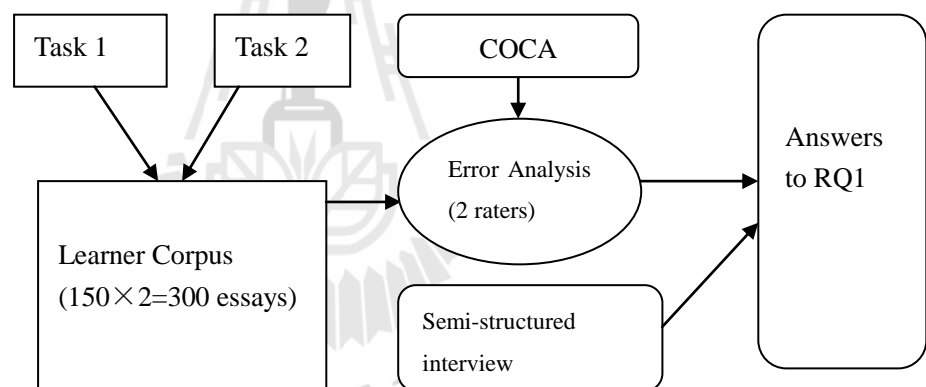
Research Questions	Instruments
1. What are typical collocation errors found in the writing of non-English majors in KU?	Two writing tasks COCA Error Analysis Semi-structured Interview
2. What are the effects of using COCA to raise students' lexical collocation awareness?	Pre-writing and Post-writing Pre-test and Posttest Instructional Treatment Questionnaire Semi-structured Interview

An overall picture of the data collection procedures of the present study is displayed in Figure 3.1. Two writing tasks were given to one hundred and fifty-one intermediate NEMs in KU who took part in CET Band-6 in December, 2010. Three hundred essays were collected as a learner corpus and compared with COCA (target language) to detect and analyze collocation errors. Afterwards, the results of error analysis were used to answer RQ1.

An intact class, 39 students, who failed in CET Band-6 in December, 2010 in KU were the participants of the quasi-experimental study, which was conducted during regular class time in an 8-week period. In the 8-week period, the subjects were required to learn lexical collocations by using COCA.

Before the instructional treatment, a pre-test and pre-writing were given to the participants. After a 6-week instruction, students took part in the posttest and post-writing to determine the effects of the collocation instruction on their collocation performance. The mean scores of the posttest were compared with the mean scores of the pre-test to measure the improvement. Correct and incorrect lexical collocations which students used in pre-writing task and post-writing task were counted to measure students' ability to use collocations. A questionnaire and a semi-structured interview were administered to participants to investigate their attitudes towards using COCA to learn collocation.

Phase 1 Error analysis



Phase 2 A quasi-experimental study

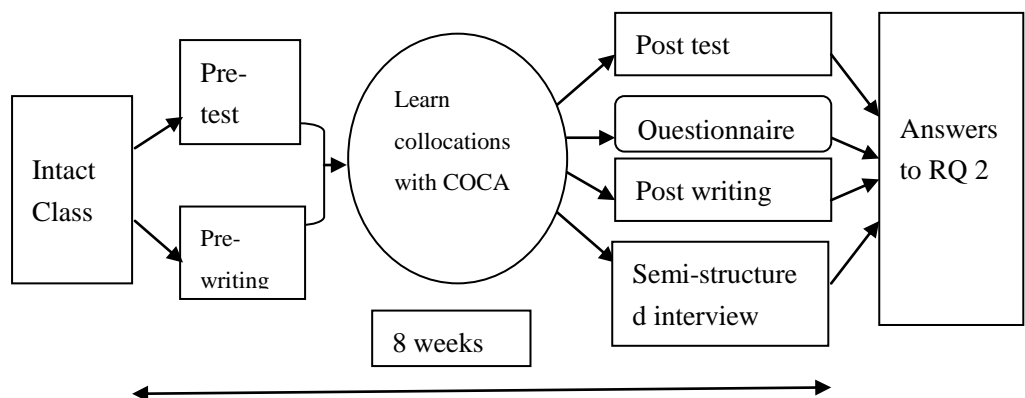


Figure 3.1 An Overview of Data Collection Procedure

3.2 Participants and Data Collection

3.2.1 Participants of the Study

Based on the literature review in Chapter 2, the focus of the present study was on investigating participants' ability to use collocations in their writing and using COCA for raising their collocation awareness. One-hundred and fifty participants were chosen in this study. The participants were intermediate level students from twenty different non-English majors in KU. These students had scored relatively high in a previous CET Band-4 before took part in the CET Band-6 in December, 2010. All of them completed two writing tasks.

An intact class of 39 students, who failed in the CET Band-6 in December, 2010 participated in the quasi-experiment. The reasons why the researcher has chosen these subjects to be the participants in the present study are as follows: 1) NEMs account for a high percentage (more than 67%) of the total number of students in KU, approximately 6800 students; 2) Compared with EMs, NEMs' English proficiency is relatively low, but it is necessary for them to master English since employers prefer students with high English proficiency; and 3) In order to become successful learners, NEMs in KU should try their best to find efficient ways to learn English.

3.2.2 Data Collection

3.2.2.1 Data from Writing Tasks & Error Analysis

Two writing tasks were administered to the participants for error analysis in the current study. They were required to write two essays on these topics: 1)

Reduce Waste on Campus; 2) How I Finance My College Education. For selecting these two topics, the researcher made a list of composition topics from CET Band-6 paper in former years. Then the researcher selected these two topics from the list of topics. The reason why these two topics were chosen was that the topics were related to the students' concerns and their college life, and they were expected to be able to perform the writing tasks with no trouble generating content and producing information. In order to get enough information, as was required in CET Band-6, the task-takers were asked to write an essay of no less than 150 words within a class period (50 minutes). To test the students' actual performance, they were required to write without the help of dictionaries and other reference books. The reason to set two compositions instead of one for this study was to reinforce the reliability of the data and to ensure understanding students' steady and actual performance in using lexical collocations.

After the students had completed the writing tasks, 300 essays (150 participants×2 essays) were collected as a learners' corpus. All lexical collocations found in the learners' corpus, as defined in Chapters 1 and 2, were underlined in terms of Benson et al's (1997) lexical collocation types (Table 2.2), and then were compared with COCA for detecting and analyzing collocation errors.

As mentioned in Chapter 2, COCA is the largest and most recent freely available corpus of English, and the only large and balanced corpus of American English. Collocation errors in the thesis refer to all the collocations which deviate from the norms

of the target language. The present study only focused on lexical collocation in which a noun, verb, adjective or adverb is combined with a word from another word class. Collocations that could not be found in COCA when MI score is at 3.0 with raw frequency at least 3 were considered collocation errors. As stated in Chapter 2, the present study took the following steps in using COCA to analyze learners' collocation errors:

Step1. Collection of a sample of learner's language and identification of errors;

This stage was a process of collecting a sample of learner's language and then identifying errors. As mentioned in the previous section, after conducting the writing tasks, 300 essays were collected as a learner corpus. All the collocations were detected and classified by two raters, separately. One rater was an experienced English teacher from the Foreign Language Institute in KU who has taught NEMs English for more than ten years. She was trained to use COCA to detect collocation errors and label them in students' compositions with error tags, and the other was the researcher herself. At the beginning, two raters read the essays twice and tried to find out the messages the subjects wish to express. Then, they underlined all of the word combinations which could be found in the participants' compositions following the working definition of lexical collocation in the present study and Benson et al's (1997) lexical collocation types (Table 2.2). In this way, grammatical collocations were excluded in the present study.

Afterward, error identification started. Identifying the participants' collocation errors was not an easy task for researchers, especially for non-native English

researchers. Therefore, in the process of detecting participants' collocation errors, COCA was used as a reference to analyze errors and provided suggestions for correction. Through utilizing COCA, it could be easier for the raters to extract examples of common authentic usages from the corpus. One important thing that they should bear in mind was that *free combinations*, such as, *close window*, *clean face*, and *dump food*, should be filtered out. Once errors were identified, the raters labeled them with error tags in terms of lexical collocation types. Lexical collocation error domains and categories are listed in table 3.2.

Table 3.2 Lexical Collocation Error Domains and Categories

	Error Domains		Error Categories
L1	Verb +noun	<L1>	Verb +noun collocation errors
L2	Adjective +noun	<L2>	Adjective +noun collocation errors
L3	Noun +verb	<L3>	Noun +verb collocation errors
L4	Quantifier + noun	<L4>	Quantifier +noun collocation errors
L5	Adverb +adjective	<L5>	Adverb +adjective collocation errors
L6	Verb +adverb	<L6>	Verb +adverb collocation errors

The following illustrates how the raters detected the participants' collocation errors in their writing samples and provided suggestions for correction by utilizing the COCA:

1. The raters found a suspicious L1 (Verb+ noun) collocation error "*They think as a student, the most important duty is to learn more knowledge from the books.*", in which the verb *learn* collocates with the noun *knowledge* in an unusual way.

2. The raters then searched the COCA (MI=3.0) with the query command *learn knowledge*, but for the search key '*learn knowledge*' no solution was yielded. Since

the components of free combinations are substitutable and their MI score are below 3.0 in COCA, to make sure whether it was a free combination, the raters queried the COCA (canceled the limitation of MI scores) with the query command again, but still no solution was displayed. So, they could decide it was a collocation error then tagged the errors. Take this sentence as an example: “*They think as a student, the most important duty is to **learn** more **knowledge**<L1> from the books.*” “*Learn knowledge*” was labeled as “<L1> L1 collocation errors” because it was a Verb+noun collocation error.

3. Next, the raters searched in COCA with the word query command *knowledge* that collocates with verb [V*] to find appropriate verbs which co-occur with the word *knowledge*. The appropriate words are *acquire* and *gain*, then find examples are shown in Table 3.3.

4. Finally, the raters gave the participants suggestions for correction: *acquire knowledge* or *gain knowledge*.

Table 3.3 Examples Extracted from COCA

-
1. *By this approach, student teachers are encouraged to **acquire** professional **knowledge** and skills by making personal efforts.*
 2. *He can not **acquire** the full **knowledge** which would make mastery of the events possible.*
 3. *She stated that constant exposure to the media was a way to **gain knowledge** and a sense of control about the war.*
 4. *Do children with visual impairments **gain** scientifically accurate **knowledge** using inquiry-based approaches?*
-

Step 2. Comparison and classification of errors

After errors were identified, the next step was to describe them properly. The description of learners' errors involved two aspects: comparison and classification. As discussed in Step 1, collocation errors could be found through utilizing COCA. In this step, the raters checked the errors by comparing the learners' errors with the reconstruction of the sentences in the target language (COCA) and check whether the two raters labeled the lexical collocation errors without making mistakes. Finally, errors were classified in terms of Benson's (1997) lexical collocation patterns, counted and calculated in terms of percentage.

Step 3. Explanation and evaluation of errors

This stage concerned discussing the reasons why NEMs in KU made unacceptable lexical collocations in their writings and evaluating errors, which was very important because some errors could reflect learners' attempts to perform the task. As illustrated in the example above, in Step 1 the raters tagged all the collocation errors in the learner corpus then in Step 2 the researcher checked all the errors. In this Step these errors were reported according to their occurrences in students' writing and calculated in terms of percentage.

3.2.2.2 Data from Pre-test, Posttest & Pre-writing, Post-writing

The Vocabulary Levels Test (VLT), originally developed by Paul Nation, and updated and validated by Schmitt and Clapham (2001), is the best-known and most widely-used vocabulary size test (receptive matching test). Matching test

provides answers for students, so it is good for low and intermediate level students. Hsu (2010) administered matching tests to investigate the effect of collocation instruction in his study. Error correction can reinforce students' collocation knowledge and can measure both receptive and productive collocation knowledge (Mallikamas and Pongpairroj, 2005).

Given what is mentioned above, in the present study, two lexical collocation tests were developed to test participants' receptive and productive knowledge of lexical collocation (see Appendix A). Each test was composed of matching and error correction parts. Matching was designed to test receptive knowledge, while error correction was designed to test productive knowledge. Each part consisted of fifteen sentences that required students to provide the best lexical collocations for the given sentences. The test was developed based on the following lexical collocations: (1) Verb + noun; (2) Adjective + noun; (3) Noun+ verb; (4) Quantifier + noun; (5) Adverb+ adjective and (6) Verb +adverb (Benson, 1997). As mentioned in Chapter 2, these six types of collocations can be taught in explicit collocation teaching. Students can use the collocation patterns to recognize collocations and collecting collocations and the researcher uses these six types of collocation to count and calculate collocation errors. The distribution of collocation types in the four tasks is listed in the Table 3.4:

Table 3.4 The Distribution of Collocation Types in Collocation Test

Tasks / Types	Matching	Error Correction
L1 Verb + noun	4	4
L2 Adjective +noun	3	3
L3 Noun+ verb	2	2
L4 Quantifier + noun	2	2
L5 Adverb+adjective	2	2
L6 Verb +adverb	2	2

Mallikamas and Pongpairoj (2005) designed two set of tasks to test the subjects' knowledge of three types of collocations: lexical, grammatical, and bound. In the tests there were 21 lexical collocations were tested. Abadi (2003) adopted a multiple-choice test of collocations with forty items among them twenty items for testing lexical collocations. Keshavarz and Salimi (2007) designed an open-ended cloze test of collocations with thirty-three items. Taking these researchers' design of collocation tests into account, a lexical collocation test with twenty to thirty items can be used to test the subjects' collocation competence. In KU, students' achievement in courses is examined through quizzes, mid-term and final exams. Usually quizzes are short tests, which are designed to be finished within 30 minutes. In the current study, a test with thirty items was designed to test students' collocation competence which could be finished within 30 minutes. According to Yang & Weir (1998), CET is of high reliability (0.90) and validity (92% of the teacher subjects believe CET reflects students' actual English proficiency levels, 86% think the test contents are reasonable). Based on CET vocabulary requirements and in order to

ensure validity and reliability, thirty lexical collocations were randomly selected as target collocation test items from ten exam papers used in CET Band-6 in five years, from 2005 to 2009. The researcher had revised items appropriately. For example, some items were made simpler and shorter, in order to make students focus on collocations and finish the test on time: “*We got to relax in the tents area, sit down and _____ a nap, tried all the cool comfy chairs.*” was revised to be “*Mama sent us to _____ a nap that afternoon after lunch.*” One exception is that the CET papers do not have matching test items. Since the subjects in the present study were intermediate students and the testing time was limited to 30 minutes, fifteen multiple choice items were revised to be matching test items.

It has been argued that the pre- and post test are not parallel forms of the same test, thus the difference between the pre- and post test score is not meaningful. However, if the difficulty level of the tests is controlled, it is theoretically acceptable to use the scores from the pre- and the post test. Paper A and Paper B in CET only reverse the order of items to ensure their content validity, reliability and difficulty level. In the present study, pre- and post test consisted of the same content except that the order of the items had been changed in the posttest. The posttest was conducted after a 6-week instruction so it was long enough for students to forget what they had done in the pre-test and the influence of pre-test on the post test was minimal. In order to prevent the participants from taking too long to answer and getting bored, the tests were designed so that students could finish it within 30

minutes. The researcher selected lexical collocations items from the ten CET Band-6 papers then made a list, after that asked a qualified and experienced English teacher who has taught NEMs in KU for more than ten years to choose target lexical collocation randomly. Considering that Verb+ noun and Adjective + noun are the central type of lexical collocation and they are tested frequently in CET, more Verb + noun and Adjective+ noun collocations are presented in the tests. The number of lexical collocations that were selected from ten CET Band-6 papers is listed in table 3.5.

Table 3.5 The Number of Lexical Collocations Selected from CET Papers

Types /Number	L1 Verb+ noun	L2 Adjective+ noun	L3 Noun+ verb	L4 Quantifier + noun	L5 Adverb+ adjective	L6 Verb + adverb
Examples	<i>tell jokes; abolish law</i>	<i>ecological balance; intellectual property</i>	<i>snow melt; baby boom</i>	<i>a pack of dogs; a set of keys</i>	<i>utterly exhausted; extremely important</i>	<i>argue heatedly; withdraw quietly</i>
Total	20	18	15	15	15	15
Select	8	6	4	4	4	4

The intact class sat for the same lexical collocation test before and after the 6-week instructional treatment. The right collocations were counted and the percentage calculated according to six lexical collocation patterns: Verb+ noun, Adjective+ noun, Noun+verb, Quantifier + noun, Adverb+ adjective, and Verb +adverb (Benson, 1997). The posttest were examined and compared with the pre-test to find out whether there is a relationship between collocation teaching and the students' knowledge of lexical collocations.

The tests were scored. The total score was 30 from 2 parts, each part was 15. One correct answer got a mark of 1 when scoring the test. In *Part 2 Error Correction*, students needed to underline each collocation error and correct it to get a mark of 1. Grammatical and spelling errors were ignored if they could not affect the raters' understanding. If a student only underlined the error or only wrote down a collocation, the answer would be considered to be wrong and he could not get a mark.

Pre- and post writing tasks were given to the intact class (39 students) before and after collocation instructional treatment for investigating students' ability to use lexical collocations. They were required to write two essays on the same topic: *Should students be allowed to live off campus in KU?* The task-takers were asked to write an essay of at least 150 words within a class period (50 minutes) without help of dictionaries and other reference books. Compared incorrect lexical collocations with correct lexical collocations the two raters found in pre-writing and post writing to measure students' improvement of using lexical collocations.

3.2.2.3 Data from Instructional Treatment & Questionnaire

According to the regulations of registering English course for NEMs in KU, NEMs who have passed CET test are considered to have achieved the requirement of the English proficiency level of graduation in KU, so they can focus on their major study and most of them do not register English courses any more. Those NEMs who have not passed CET test are divided into classes to learn English

courses. An intact class of NEMs in KU took part in collocation instructional treatment. As mentioned before, they were participants took part in CET Band-6 which was held in December, 2010. They had passed CET Band-4 and failed in CET Band-6.

Hsu (2010) conducted a research study to examine the effects of a 9-week explicit collocation instruction on Taiwanese college English majors' reading comprehension and vocabulary learning. Only in Week 1, Week 4, and Week 7 were collocation instructions given and each class lasted one hour. In other weeks, 20-minute vocabulary tests were implemented. Lien (2003) was the first to explore the effects of collocation instruction on reading comprehension. She designed and carried out a quasi-experimental study at a national university in central Taiwan over a four-week period. Based on previous studies, the researcher took the following factors into consideration for designing collocation instruction.

1. Collocations are too numerous for all of them to be taught in classroom.

In Collins' COBUILD English Words in Use, more than 100,000 collocations are given, covering the core vocabulary of English. There are more than 70,000 collocations listed under about 14,000 entries in BBI Combinatory Dictionary of English (Bahns, 1993).

2. The process of acquiring collocation is a time-consuming process. At present, the best conservative rule of thumb that we have is that up to a vocabulary size of around 20,000 word families, we should expect that native speakers will add

roughly 1000 word families a year to their vocabulary size (Goulden, Nation and Read, 1990). Native speakers usually have their ability to dispose thousands of words. They have the power to predict the meaning when two or more words are combined together and tend to use a great number of words and put them together in various ways according to their spoken and written communication needs. However, it is hard for EFL/ESL learners to reach this level although they have been learning English for more than 10 years.

3. While taking collocation teaching into consideration, we need to bear in mind that not all of the collocations have to be learned. Focus needs to be placed on building students' consciousness of how words work in combination with one another, so they can continue developing their collocation competence after class.

In the present study, an 8-week quasi-experimental study was employed. As mentioned in Chapter 2, collocation instruction focused on raising participants' collocation awareness through using COCA. In the first week, pre-test and pre-writing were given to the participants (80 minutes). The following six weeks were used for learning the six types of lexical collocations via using COCA (one class period for one week). In the last week, posttest, post -writing, and questionnaire were administered to the students (8 weeks in all).

Before the instructional treatment, students registered COCA as free users. Focusing on using COCA to raise students' collocation awareness, the researcher selected several different content words in different parts of speech from students'

text book (New Horizon College English Book 3) as examples to introduce different type of lexical collocations in each class. Also, some passages and sentences were selected from the text book to design exercises for using COCA to raise students' lexical collocation awareness. Every class followed the same teaching steps and time allotment. During the instruction students watched videos of COCA tutorial and introduction of collocation. Students were required to do fill-in-blank, matching, collocation collection, and error correction exercises these were designed according to Benson's lexical collocation types (one type for one class). Lesson plan see: Appendix A. The arrangement of instruction is shown in Table 3.6.

Table 3.6 Treatment Arrangement in the Present Study

Intact Class-----Experiment Group (EG)	
Week 1	Pre-test 30 minutes; Pre-writing 50 minutes
Week 2	Learn Verb + noun collocations with COCA (50 minutes)
Week 3	Learn Adjective + noun collocations with COCA (50 minutes)
Week 4	Learn Noun + verb collocations with COCA (50 minutes)
Week 5	Learn Quantifier + noun collocations with COCA (50 minutes)
Week 6	Learn Adverb + adjective collocations with COCA (50 minutes)
Week 7	Learn Verb + adverb collocations with COCA (50 minutes)
Week 8	Posttest 30 minutes; Post-writing 50 minutes; Semi-structured Interview 30 minutes Questionnaire 30 minutes

Questionnaires can be designed and used to collect vast quantities of data from a large number of subjects. They can also be used to elicit information about many different kinds of issues, such as language uses, preferred learning styles, attitudes and beliefs. They have a number of benefits over other forms of data

collection: they are relatively easy to prepare, inexpensive to administer, and they obtain information that is relatively easy to tabulate and analyze (Richards, 2001). The questionnaire in this study was used to elicit information about subjects' attitudes towards using COCA to learn lexical collocation, and their English proficiency and recommendations. Thus, the questionnaire consisted of three parts: 1) Background information; 2) Attitudes toward using COCA to learn lexical collocation; and 3) English proficiency and recommendation. All of the questionnaire items were translated into Chinese to avoid misunderstanding and confusion (see Appendix C).

In order to get more insightful information from participants and to know the effect of collocation treatment, the questionnaire was conducted only for the students who had participated in the collocation instruction. The designing of the questionnaire was based on lexical collocation types (Benson, 1997) and using COCA.

3.2.2.4 Data from interview

In the present study, a semi-structured interview (See Appendix D) was conducted to elicit more information about students' attitudes towards making collocation errors and using COCA to learn collocations. Robson (2002) stated that an interview is "a conversation initiated by the interviewer for the specific purpose of obtaining research-relevant information and focused by him or her on content specified by research objectives of systematic description, prediction or explanation" (p. 97). A face-to-face interview is a "shortcut" in seeking answers to research questions because it offers the researcher the opportunity to ask participants directly about what is going on (Robson,

2002). Interviews can be classified into three categories i.e. unstructured interview, semi-structured interview, and structured interview (Nunan, 2001).

The reason why the researcher chooses a semi-structured interview lies in the fact that semi-structured interviews are flexible and it may be the most popular among the three categories. The semi-structured interview gives the interviewee a degree of power and control over the course of the interview. Since the purpose of this semi-structured interview is to elicit more insightful information about making collocation errors and using COCA to learn collocation, not test students' English proficiency. For the sake of convenience, both English and Chinese were used for better understanding. The researcher contacted and interviewed them by using the video tool of Tencent QQ, which is a popular free instant messaging computer program in China. The interviews were tape-recorded and transcribed for data analysis.

Twenty interviewees were selected for the interview. They made collocation errors in five different types of collocation i.e. L1 Verb + noun, L2 Adjective + noun, L4 Quantifier + noun, L5 Adverb + adjective, and L6 Verb + adverb. Four students were selected by they made the four most frequent collocation errors in each type, in total twenty interviewees. These students were asked three questions about why they made the four most frequent errors in each type of collocations. Among the twenty interviewees, there were six students came from the intact class, who were asked two more questions about their attitudes towards using COCA to learn collocations (See: Appendix D).

3.3 Data Analysis

This section discusses the methods for data analysis. After collecting the data, the researcher analyzes the data by Statistical Package for the Social Science 17.0 for Windows (SPSS). The method of data analyses in the study involves both quantitative and qualitative analysis.

3.3.1 Qualitative Analysis

Data collected from students' writing tasks were analyzed by error analysis qualitatively to see the weaknesses of the students' use of collocation. Furthermore, data collected from students' questionnaires and interviews were analyzed qualitatively to describe the subjects' answers to open-ended questions about their opinions towards using COCA to learn collocation, English proficiency, and recommendations to learn collocations.

3.3.2 Quantitative Analysis

Paired sample t-test was used to compare the participants' mean scores on the pre- and post test. The purpose was to see whether there were significant statistical differences between students' pre- and post test scores, thus to decide the effects on improving collocation performances of students.

Data obtained from the questionnaire were input into SPSS 17.0 to find out the results of individual ideas towards difficulty levels of lexical collocation types and attitudes of using COCA to learn collocation.

3.4 The Pilot Study

According to Lancaster, Dodd, and Williamson (2004), a pilot study or feasibility study is designed to test the logistics and gather information prior to a large study in order to improve the latter's quality and efficiency. A pilot study may address a number of logistical issues. The pilot study is related to the examination and development of the instruments for Research Questions 1 and 2 in the present study. The purposes of this pilot study in the present study are: 1) to check the weaknesses of students' use of lexical collocations; 2) to check the practicability of error analysis procedure; 3) to check reliability and validity of collocation test and questionnaire; 4) to check students' correct operation of COCA; and 5) to check the instruction given to students is comprehensible.

3.4.1 Participants and Procedures

Fifty NEMs from KU participated in two writing tasks. They were intermediate students from eight different non-English majors. These students had scored relatively high in a previous CET Band-4 before they took part in the CET Band-6 in December, 2010. The two writing tasks were conducted on 20th, May, 2011 and lasted two class periods (50×2=100 minutes). The participants were required to write two essays on these topics: 1) Reduce Waste on Campus; 2) How I Finance My College Education. These two tasks were topics familiar to the participants. In order to get enough information, as is required in CET Band-6, the participants were asked to write both essays in at least 150 words within two class period (50×2=100

minutes). To test the students' actual performance, they were required to write without the help of dictionaries and other reference books when doing the tasks. Data collected from students' writing tasks were analyzed by two raters following the error analysis steps, as stated in Chapter 2, to analyze collocation errors and see the weaknesses of the students' use of collocations.

Ten NEMs from KU participated in the experiment. They were intermediate students from three different non-English majors. These students were participants in the CET Band-6 in December, 2010 and were required to take the pre-test which was conducted on October, 17th, 2011 and lasted 30 minutes. As mentioned previously, mean scores from pre-test were compared with posttest mean scores to decide the effects on students' improvement of collocation performances. The instruction began three days after the pre-test, on October, 21th, 2011. During the instruction (60 minutes), the participants were taught to register COCA (10 minutes) and learned how to use COCA to do collocation exercises (50 minutes). Posttest and questionnaire were conducted ten days after the pre-test, on October, 27th, 2011 and lasted 60 minutes (30 minutes for posttest, 30 minutes for questionnaire).

3.4.2 Results of the Pilot Study

3.4.2.1 Results of Error Analysis

Before the writing samples were collected, the demarcations between lexical collocations and non-lexical collocations were made. As mentioned in Chapter 2, in the present study, only word combinations with MI scores higher than

3.0 and raw frequencies higher than 3 in COCA are classified as collocations. In the process of data collection, collocation errors are identified, counted, and percentage calculated in terms of Benson's (1997) lexical collocation patterns (Table 2.2). Error occurrences and their percentage are reported in terms of error tags.

After conducting the writing tasks, as mentioned in 3.4.1, one hundred essays (50×2) were collected, which constituted a learner corpus. The learner corpus included 19,140 words. The average length of the writing was 191 words. After reading the participants' compositions twice, the researcher began to detect lexical collocations. All lexical collocations found in the learners corpus as defined in Chapters 1 and 2 were underlined in terms of Benson et al's (1997) lexical collocation types (Table 2.2). As mentioned in Chapter 2, in the present study the collocation is limited to lexical collocations, so grammatical collocations, idioms and free combinations that appeared in the writing samples were sorted out in this research.

As discussed in 3.2.3, the present study using COCA to analyze learners' collocation errors followed three steps. In Step 1, the two raters underlined all the word combinations in the learner corpus separately. Then, both queried all the underlined combinations in COCA and preliminarily judged which word combinations were collocation errors. In Step 2, the researchers compared collocation errors from the learner corpus with the reconstructed sentences in COCA to check whether the two raters' judgments and classifications of errors were correct and

whether they tagged errors in the learner corpus without making mistakes. At the end, they found 214 lexical collocations in the subjects' writing, which included 106 error-free collocations and 108 incorrect ones. The findings are shown in are shown in Table 3.7.

Table 3.7 Number and Percentage of Correct Collocations and Collocation

Errors in Learner Corpus					
Type	Total No.	Correct Collocations Number (%)	Examples	Collocation Errors Number (%)	Examples
L1—Verb+ noun	107	52 (48.6%)	<i>solve problem;</i> <i>attend college;</i>	55 (51.4%)	<i>Learn knowledge;</i> <i>fit the demand;</i>
L2—Adjective +noun	64	35 (55%)	<i>negative effects;</i> <i>serious problem;</i>	29 (45%)	<i>smooth English;</i> <i>social sign;</i>
L3—Noun+ verb	0	...	0	...	0
L4—Quantifier + noun	5	1 (20%)	<i>A pair of jeans</i>	4 (80%)	<i>a line of hope ;</i> <i>a cloud of wind ;</i>
L5—Adverb+ adjective	12	8 (66%)	<i>extremely high;</i> <i>highly educated;</i>	4 (33%)	<i>definitely value;</i> <i>intensely fresh;</i>
L6—Verb +adverb	26	10 (38%)	<i>affect deeply;</i> <i>reading widely;</i>	16 (62%)	<i>manage reasonably;</i> <i>arrange properly;</i>
In Total	214	106 (49.5%)		108 (50.5%)	

Some collocations repeatedly appeared in the subjects' writing, such as: *learn knowledge*, *daily life*, and *add parents' burden*. In this case, they were only counted once. Some collocations such as *practice their thrifty conscious* (raise their thrifty consciousness) were counted twice, and *practice conscious* and *thrifty conscious* were considered two errors.

As the data shown in Table 3.7, there was no L3 Noun+ verb lexical collocation in this learner corpus. This is probably because L3 lexical collocations, such as *lion roar*, *clock chime*, and *bell ring*, were rare for them to use in argumentative writings. Errors amounted to a large percentage in L4 Quantifier + noun lexical collocations (80%), but it does not mean that this type of collocations is the most difficult one. Actually, the compositions indicated that participants seldom use this L4 pattern of collocations. The participants employed only five L4 lexical collocations in their writings. The most frequent errors were related to the use of L1 Verb +noun collocations. The participants made 55 errors in this type. Figure 3.2 below displays the distribution of lexical collocation error domains found in the present learner corpus.

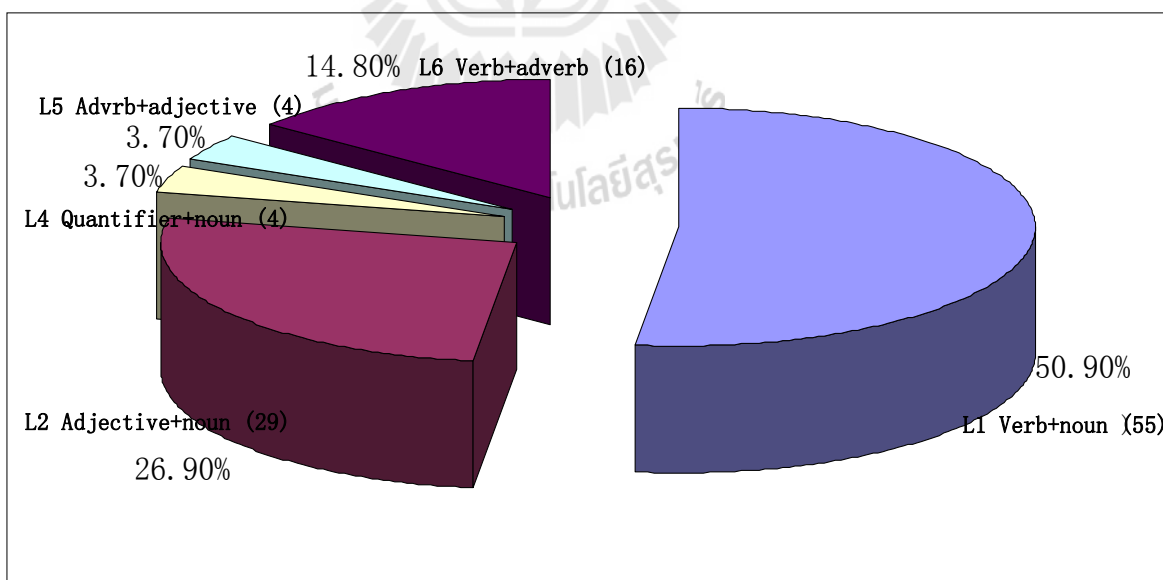


Figure 3.2 Distribution of Lexical Collocation Errors

Of all the collocation error types, the most frequent one was the collocation errors in L1 Verb + noun, totally 55 (50.90%). On one hand, this is because L1 Verb+noun collocations are used most frequently in this learner corpus. On the other hand, a verb in a collocation has a restricted sense, which makes its correct use more difficult when learners cannot fully distinguish subtle differences among verb candidates, for instance, *tell jokes*, *tell a story* but not *say jokes*, *say a story*. Therefore, Chinese EFL learners have more trouble in choosing a proper verb in collocations (Chen, 2002). The second most frequently used type is L2 Adjective + noun collocations, totally 29 (26.90 %). This is because adjectives were used second most frequently in their compositions and probably it was hard for students to distinguish and select appropriate adjectives to express their meanings.

Collocation errors were detected and classified by two raters respectively. Cronbach alpha reliability test indicated that the inter-rater reliability ($\alpha=0.854$) was quite satisfactory. In order to know how frequently a certain pattern of collocation errors appeared in the learner corpus, the researchers listed the findings of lexical collocation error occurrences and the percentage in Table 3.8.

Table 3.8 Distribution of Lexical Collocation Error Domains

Tag	Number of Errors	Number of Occurrences	Percent (%)
L1	55	88	57.9
L2	29	37	24.3
L4	4	5	3.3
L5	4	5	3.3
L6	16	17	11.2
Total	108	152	

Note: Error tags as shown in Table 3.2

From Table 3.8, a total number of 88 tokens were found, representing 55 <L1> errors, which accounted for 57.9% of the total tokens. That is, 11 errors occurred more than once. There were 29 <L2> errors realized by 37 tokens in <L2> type, which accounted for 24.3 %. That is, 7 errors were repeated. When comparing <L5> Adverb+ adjective with <L6> Verb+ adverb, participants made more errors in <L6> pattern. That is, 16 <L6> errors were realized by 17 tokens, which accounted for 11.2% and ranked the third most frequent one in the learner corpus. <L5> and <L4> patterns of errors were the least frequent in the learner corpus with 5 tokens representing 4 errors each, which accounted for 3.3 %, respectively.

All error types that occurred in the writings are discussed below in details:

Some students translated Chinese into English literally according to the Chinese linguistic conventions. These expressions are understandable in Chinese but are not acceptable in English. Here are a few of such errors:

*We can't **fit the demand**<L1> (to meet the demand) of finding a job.*
*I don't want to **add** my parents' **burden**<L1> (impose a burden) and make them work hard for me.*
*I still have **a line of hope** <L4> (a glimmer of hope) I can earn money to pay the tuition fee by my own, but parents didn't think so.*
*Many students **open the lights** <L1> (turn on the lights) even though they go out to play.*
*We should try our best to make full use of study time, instead of wasting it reading more widely and **studying more knowledge** <L1> (acquire/gain knowledge).*
*The most important is that I can **learn** all kinds of **knowledge**<L1> (acquire/gain knowledge).*

Students who had learnt that some verbs + ing/ed function as adjectives and knew some collocations such as *a missing boy*, *a knowing smile*, probably considered verb + ing equaled adjectives, hence errors occur, which are illustrated by these examples:

*In order to have an **attracting looking**<L2> (an attractive looking) they spend much money.*
*Besides it can also create a **healthing environment** <L2> (healthy environment) for us.*

Owing to the lack of collocation knowledge, some verbs such as *take*, *make*, *do* have a very wide applicability but have little meaning on their own. Some students might have thought these verbs could replace one another. And some learners use synonyms to replace note words or collocates. Therefore, the participants made errors such as:

People can **make advantage**<L1> (take advantage) of saving time.
 We should **take a profit** <L1> (make a profit) on doing part-time job.
 We **do plans**<L1> (make plans) for summer holidays.
 We also need to **treasure every moment** <L1> (cherish every moment) in our daily life.
 We should care for the environment and try to **less unnecessary waste** <L1> (reduce waste).
 I think that it is **definitely value** <L5> (definitely worth) the risk to take the part-time job.
 The social support **provides many chances**<L1> (provide opportunities) to student.
 Few of them can speak **smooth English**<L2> (fluent English).
 I believe that we can reduce the waste on campus with our **insistent efforts**<L2> (ceaseless efforts).

Some students were used to learning words in isolation. They only understood the basic meanings of the word but did not know what words it would go with. It is because they were not able to produce the right collocation or they just put words together randomly:

The most important is to **plant the conscious** <L1> (raise consciousness) of reducing waste in daily life.
 It's easy to find that most students try to **practice their thrifty conscious** <L1> (raise their thrifty consciousness).
 It takes lifetime to **prove promise** <L1> (to fulfill promise).
 If we have a good habit of save energy and foods, **keeping the environment balance** <L2> (keeping the ecological balance), we would have a lovely planet.
 I stay in classroom when my classmates like **a cloud of wind**<L4> (a gust of wind) to go to supermarket to buy goods after class.
 Waste resource is a main reason of causing **strongly polluted**<L5> (badly / seriously /heavily polluted) environment.
 Time like arrow, we should grasp it and **manage reasonably** <L6> (manage effectively or properly).

As mentioned in the previous parts, for error analysis, the inter-rater reliability ($a=0.854$) was quite high. Actually, in the pilot study, the two raters arrived at an

agreement when using COCA to detect collocation errors. For instance, they decided “span” for a query key (4 words on the left and 4 words on the right of a “node” word) when they detected what words go with a given word. However, some practical limitations occurred. The two raters were free users of COCA. Sometimes the corpus was unstable, and query results could not be yielded. They needed to log out and then log in again. Also, the number of queries was limited for free users each day, so each rater had to register two accounts to avoid slowness in conducting searches. Since COCA can be used as an assistant tool to raise students’ collocation awareness in the main study, the researchers take all the limitations into consideration.

3.4.2.2 Piloting Instructional Treatment

Before the instructional treatment, the researcher spent ten minutes to teach students how to register COCA. For the sake of convenience, all the ten participants used their QQ mailbox and password to register COCA as free users because all the students had QQ mailbox.

The content of piloting instructional treatment was “Using COCA to learn L1 Verb+ noun Collocations” which were conducted in week 2 in the main study. Focusing on using COCA to raise students’ collocation awareness, the researcher selected two verbs (*relax* and *entertain*) from students’ text book (New Horizon College English Book 3) as examples to introduce L1 Verb+ noun lexical collocations in the class. Also, two passages and three sentences were selected to

design exercises for using COCA. Students were required to do fill-in-blank, matching, collocation collection, and error correction exercises which were designed according to Benson's (1997) lexical collocation types.

In order to make the instruction given to the students in the main study comprehensible and every class finish in time, the researcher adjusted the time allotment of the lessons and added one more teaching step (Step 7) in lesson plan after the instructional treatment. Furthermore, in the piloting of the treatment, only ten minutes for students to register COCA was not enough. Therefore, in the main study before the instructional treatment, the researcher should give more time to students and make sure all the students have registered COCA already before the instruction. The lesson plan is shown in Appendix A.

3.4.2.3 Comparison between the Collocation Pre-test and Posttest

To check whether the items of the collocation test could measure what they were designed for, the test together with the evaluation form for content validity check were sent to three experts who were all professors and were academically qualified in China. The experts checked the evaluation form by using Item-Objective Congruence Index (IOC) as a validation method for the relevancy of the content and the objective of the test. The researcher adjusted the test according to the experts' suggestions and checked the results of IOC index for each item by item analysis (IAS). According to Booncherd (1974), the acceptable value should be higher or equal to 0.5 (≥ 0.5), the content validity value of current collocation test was 0.843 which

indicated the test was acceptable. The results indicated that there were 5 unacceptable items in the 30 items of the test, the researcher improved and corrected unacceptable items according to the 3 experts' opinions and suggestions. The same two raters scored the test, and then Cronbach alpha reliability test of scoring the tests by the two raters indicated that the inter-rater reliability is $\alpha=0.89$ which is higher than the cut-off point 0.7. That means the agreements of scoring the tests by the two raters are significant.

Ten participants received collocation tests. The posttest was administered and the scores were compared with those of the pre-test to determine the effects of the collocation instructional treatment. Both tests were scored according to the marking criteria presented in 3.2.2.2 of Chapter 3. Compared with the pre-test, the overall results were negative because the majority of students' performance did not improve in the posttest.

The results from Paired samples t-test showed that there is no significant difference between the pre-test and posttest scores because the p value is over 0.05 ($p = 0.535 > 0.05$), and the mean scores of the pre-test and the posttest are nearly the same (11.15 and 11.72). A positive difference of 0.57 in the mean score indicates that there was a slight improvement in the posttest. It suggests that students made no significant improvement on their collocation performance after the instructional treatment. After piloting the collocation test, the researcher modified two items in the test (see Appendix B).

3.4.2.4 Results of the Questionnaire

In order to check the content validity of all the items and questions in students' questionnaire, the same three experts were invited to validate and check the language use for each item. According to Booncherd (1974), the acceptable value should be higher or equal to 0.5 (≥ 0.5). The content validity value of the current questionnaire was 0.867 which indicates the questionnaire was acceptable. After revising some parts according to the experts' suggestions, the items and questions were suitable for the purpose of the questionnaire in the present study (see Appendix C).

A 5-point Likert-scale questionnaire that ranged from "strongly disagree" to "strongly agree" was employed. After the instructional treatment, the students were required to answer the questionnaire. Students' responses to the questionnaires were coded and input into SPSS 17.0 for statistical analysis. The responses are described in Table 3.9.

Table 3.9 Students' Responses on the Likert-scale Questionnaires (N=10)

Item	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1 I think COCA is useful in searching for English collocations. (我认为 COCA 对查询英语词语搭配很有用)	0%	10%	20%	20%	50%
2 I think I will use COCA as an assistant tool to learn English collocations in the future. (我想将来我将使用 COCA 作为辅助工具来学习词语搭配)	0%	20%	20%	60%	0%
3 I think COCA is useful in helping me identify collocation errors in English. (我认为 COCA 能帮助我识别英语词语搭配的错误)	10%	20%	20%	30%	20%
4 I enjoy using COCA as support for learning English collocations. (我很喜欢用 COCA 来学习英语词语搭配)	10%	0%	20%	70%	0%
5 I think COCA is user-friendly. (我认为 COCA 很便于操作)	0%	10%	30%	20%	40%
6 I think I will recommend using COCA as support for learning English collocations to other students. (我想我将推荐 COCA 给其他同学学习英语词语搭配)	0%	10%	20%	70%	0%

Table 3.9 shows the responses from students' questionnaires. On one hand, the majority of the students think that COCA is useful in identifying collocation errors and searching for English collocations, and they will not only use COCA in the future but also will recommend COCA to other students to learn English. For Items 1, 6, and 4, the percentage of students who agreed that *I think COCA is useful in searching for English collocations*, *I think I will recommend using COCA as support for learning English collocations to other students*, and *I enjoy using COCA as support for learning English collocations* are 70%. For Items 2 and 5, the percentage of students who agreed that *I think I will use COCA as an assistant tool to learn English*

collocations in the future and *I think COCA is user-friendly*, are 60%. For Item 3, the percentage of students who agreed that *I think COCA is useful in helping me identify collocation errors in English* is 50% (20+30). On the other hand, a negative aspect is mentioned and should be considered carefully, which is that some students considered that COCA is not user-friendly, and the percentage was 30%.

In the part of the open-ended questions from students' questionnaire, the students reported that they have been learning English for more than ten years. A majority of students reported that they had received teachers' in-class support in using online resources, such as, Baidu Video, and New Orient online English courses. Some students recommended their ways to learn collocations to other students, such as translating sentences, using notebook, and Association Memorizing Methods.

In order to determine the reliability of the questionnaire, Cronbach's Alpha (α) was used to check the internal consistency of the questionnaire items by analyzing the data from the pilot study. According to DeVellis (2003), acceptable reliability of the questionnaire is reached if the alpha is no less than 0.70 ($\alpha \geq 0.70$). The reliability from the pilot study was 0.83, which was higher than 0.70 ($\alpha=0.83 \geq 0.70$). Therefore, all of the items in students' questionnaires in the present study were considered reliable and could be used in the main study (see Appendix C).

3.5 Conclusion and Implications for the Main Study

There are two main limitations of the pilot study. First, because of the short period of the instruction, the researcher only taught one type of lexical collocations to check time allotment and whether the teaching contents are suitable or comprehensible. As a result, there was no chance for the researcher to get further informative data. In the main study, all lexical collocation types were taught and the instrument was applied to get further informative data during the 8-week instruction. Second, students' posttest scores in the pilot study might be influenced by their pre-test ones because of the short period of time between these two tests (only ten days). They might still remember what they have done in their pre-test and they might try to perform better in their posttest. However, in the main study, which takes 8 weeks of the quasi-experimental study, this problem may not affect students' performance on the posttest.

Several implications from the pilot study can be summarized. First, Benson's (1997) lexical collocation types are practicable to classify lexical collocations in this learner corpus; Second, COCA is not only useful to detect and identify lexical collocation errors but also practicable to query lexical collocations for data-driven learning; Third, the error analysis procedure is workable to analyze lexical collocation errors; Fourth, the implementation of CET papers to design pre-test and posttest is also suitable and exercisable. Last, the use of both English and Chinese in students' questionnaire is appropriate, because students have no difficulty understanding each

item. Therefore, collocation test, questionnaire, Benson's (1997) lexical collocation types, COCA, and the error analysis procedure will continue to be used in the main study.

3.6 Summary of Chapter 3

This Chapter discusses the research methodology employed in the present study. The methods of data analyses in the study involve both quantitative and qualitative analyses. Error analysis was administered to examine the weaknesses of using lexical collocations in NEMs' writing in KU qualitatively. An eight-week quasi-experimental study was implemented to raise students' collocation awareness. SPSS 17.0 was used to analyze the results of individual ideas and attitudes from student questionnaire. Paired samples t-test was employed to analyze students' scores on the collocation pre-test and posttest. In the next chapter, the research results and findings will be presented in details.

CHAPTER 4

RESULTS AND DISCUSSION

In this chapter, the results of the main study are presented and provided discussion on those results. Section 4.1 presented the results of error analysis in the learner corpus which offers answers to Research Question 1: What are the typical lexical collocation errors found in the writing of non-English major EFL learners? For Research Question 2: what are the effects of using COCA to raise students' lexical collocation awareness? In Section 4.2, comparison of error analysis results from the pre- and post writings of 39 student participants who took part in the instructional treatment using COCA to raise their lexical collocation awareness is discussed to find out changes or improvement in their use of collocation in their writing. Also, results from the pre- and post collocation tests using SPSS 17.0 are explained to reveal these students' changes in their performance. Finally, findings from the questionnaire and interview to uncover the students' attitudes towards the use of COCA to learn collocations are summarized. In Section 4.3, a summary of Chapter 4 is presented.

4.1 Typical Lexical Collocation Errors in the Learner Corpus

As discussed in the previous Chapter, 151 intermediate level NEMs in KU took part in two writing tasks in the present study. In the main study, the researcher

probed deeply into the learner corpus. Two raters followed the error analysis procedure, which was presented in Chapter 3, using COCA to analyze learners' collocation errors in the learner corpus (300 essays, 57,900 words, with the average length of writing being 193 words). One student whose two compositions were short, only 65 and 76 words, respectively, and their content was not related to the topics so these compositions were excluded from the present study.

Collocation errors were detected and classified by two raters separately. As shown in Chapter 3, Cronbach alpha reliability test of the collocation identification by the two raters indicated that the inter-rater reliability is $\alpha=0.854$ which is higher than the cut-off point 0.7. That means the agreements of identification results by the two raters are reliable.

In Step 1, the two raters underlined all the word combinations in the learner corpus separately. Then, both queried all the underlined combinations in COCA and preliminarily judged which word combinations were collocation errors. In Step 2, the researchers compared collocation errors from the learner corpus with the reconstructed sentences in COCA to check whether the two raters' judgment and classification of errors were correct and whether they tagged errors in learner corpus without making mistakes. At the end, the raters found 310 lexical collocations in the learner corpus, which included 143 correct collocations and 167 incorrect collocations. The findings are shown in Table 4.1.

Table 4.1 Correct and Incorrect Collocations in the Two Writing Tasks

Type	Total No.	Correct collocation Number(%)	Examples	Incorrect collocation Number (%)	Examples
L1—Verb+ noun	162	77 (47.5%)	<i>solve problem;</i> <i>make contribution;</i>	85 (52.5%)	<i>learn knowledge;</i> <i>plant conscious;</i>
L2—Adjective +noun	79	33 (41.8%)	<i>negative effects;</i> <i>remote areas;</i>	46 (58.2%)	<i>smooth English;</i> <i>smeary environment;</i>
L4—Quantifier + noun	6	2 (33.3%)	<i>a pair of jeans</i> <i>a fit of anger</i>	4 (66.6%)	<i>a line of hope;</i> <i>a bundle of flowers;</i>
L5—Adverb +adjective	17	9 (52.9%)	<i>extremely high;</i> <i>highly educated;</i>	8 (47.1%)	<i>definitely value;</i> <i>strongly polluted;</i>
L6—Verb +adverb	46	22 (47.8%)	<i>affect deeply;</i> <i>reading widely;</i>	24 (52.2%)	<i>manage reasonably;</i> <i>arrange properly;</i>
In Total	310	143 (46.1%)		167 (53.9%)	

As shown in Table 4.1, the results of error analysis in the main study were similar to those in the pilot study. The most frequent error type was L1 Verb + noun collocations. The participants made 90 errors of this type of lexical collocations. The second most frequent error type was L2 Adjective + noun collocations, i.e. 46 errors. Thus, the main difficulty for ENMs in KU to use lexical collocation was in selecting verbs and adjectives to express themselves. There was no L3 Noun+ verb lexical collocation that the subjects produced in this learner corpus. This is probably because L3 lexical collocations, such as *loudspeaker screech*, *thunder rumble*, and *glasses clink*, were hard for intermediate level NEMs in KU to master and use. Errors amounted to a large percentage in L4 Quantifier + noun lexical collocations, i.e. 66.6%, but the analysis revealed that the participants

seldom used L4 pattern of collocations because there were only 6 L4 lexical collocations in the corpus. Therefore, one of weaknesses for NEMs in KU to use collocation was in choosing what quantifier to go with a noun. Figure 4.1 below displays the distribution of lexical collocation error domains found in the present learner corpus.

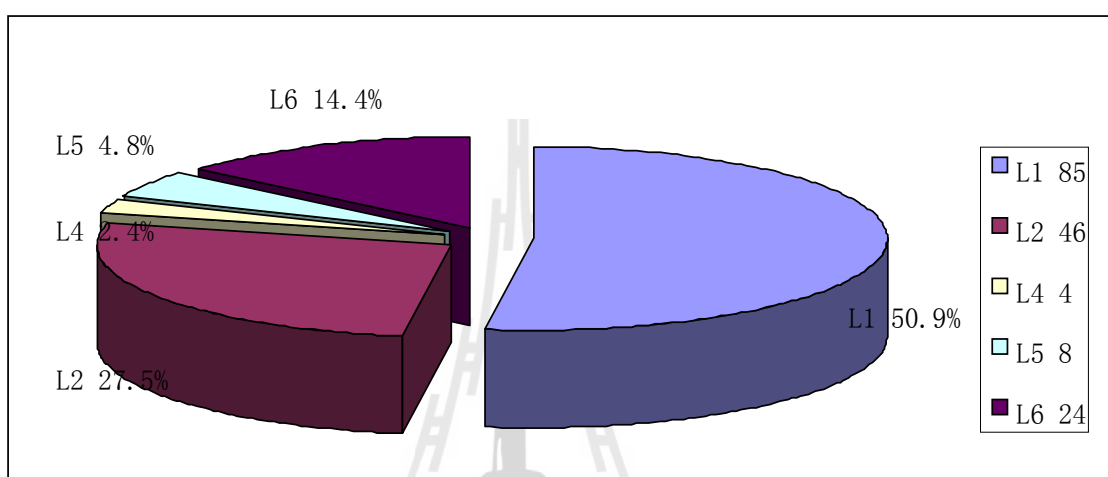


Figure 4.1 Distribution of Lexical Collocation Errors

As shown in Figure 4.1, of all the collocation error types, the most frequent one was the collocation errors in L1 Verb+ noun i.e. 85 (50.9%). The second most frequent type was L2 Adjective+ noun collocations i.e. 46 (27.5 %), and then L6 Verb+ adverb i.e. 24 (14.4%). Probably, a verb in a collocation has a restricted sense, which makes its correct use more difficult when learners cannot fully distinguish subtle differences among verb candidates. Also, it is hard for students to distinguish and select appropriate adjectives to express their meanings.

In order to know how frequently a certain pattern of collocation errors appeared in the learner corpus, the researcher listed the findings of lexical collocation error

occurrences and percentage in Table 4.2.

Table 4.2 Distribution of Lexical Collocation Error Domains

Tag	Number of Errors	Number of Occurrences (Tokens)	Percent (%)
L1	85	152	48.4
L2	46	91	29.0
L4	4	7	2.2
L5	8	26	8.3
L6	24	38	12.1
Total	167	314	100%

Note: Error tags as shown in Table 3.2

From Table 4.2, a total number of 152 tokens were found representing 85 L1 errors, which accounted for 48.4% of the total tokens. That is, 23 errors occurred more than once. There were 46 L2 errors realized by 91 tokens in L2 type, which accounted for 29.0%. That is, 14 errors were repeated. Compared L5 Adverb+ adjective (26 tokens found representing 8 L5 errors) with L6 Verb+ adverb (38 tokens representing 24 L6 errors), participants made more errors in L6 pattern of errors, , which accounted for 11.9% and ranked the third most frequent one in the learner corpus. L4 pattern of errors was the least frequent errors in the learner corpus (7 tokens presenting 4 errors, which accounted for 2.2 %).

The results of error analysis have shown that the most frequent error type was collocations with verb as nodes, i.e. L1 Verb+ noun. There were 90 L1 Verb+ noun errors (accounted for 52.3%) in the learner corpus. The second most frequent type was collocations with adjectives as nodes, i.e. L2 Adjective+ noun. There were 46

L2 Adjective+ noun errors (26.7 %) in the learner corpus, and misuses of quantifiers were also found to be common in the corpus because the learners only produced 6 L4 Quantifier+ noun pattern of collocations, of which 4 were errors. Compared L5 Adverb+ adjective (8 errors, accounted for 4.7%) with L6 Verb+ adverb (24 errors, accounted for 14.0%), students had more trouble in choosing an adverb to go with a verb than choosing an adverb to go with an adjective.

As mentioned in Chapter 3, twenty students took part in the semi-structured interview. According to their report, the following reasons can be used to explain why they made collocation errors. First of all, they translated Chinese into English, for instance, *learn knowledge*, *study knowledge*. Secondly, students who made L4 Quantifier+ noun collocation errors reported that it was hard for them to choose a quantifier to go with a noun. Thirdly, it was difficult to master the usage of *make*, *do* and see the differences between synonyms. Lastly, to learn which verb can be added an *ing* to function as an adjective was difficult too. For the reason it is difficult for them to master collocations, the participants reported that they lack vocabulary, so they have no sense which word can go with which word. For how to correct the errors, they reported that they can use the dictionary, search online, make notes, and ask teachers and classmates for help.

4.2 Discussion

The findings of error analysis in the present study suggested that, first of all,

majority of the collocation errors are in the two categories of L1 Verb+ noun and L2 Adjective+ noun. Students have problems with choosing words and expressions to make their sound natural, especially in choosing verbs and adjectives. The findings are consistent with the findings of similar studies in literature (Chen, 2002 and Fan, 2009). According to Chen (2002), learners cannot fully distinguish subtle differences among verb and adjective candidates, even English high achievers in college have problems with choosing words to express their ideas. In terms of the uses of collocations with adjectives as nodes, Fan (2009) found that evidence of deviant collocational use resulting from the learners' confusion with the English language. For example, some errors occurred in Hong Kong learners' writing, such as *curve hair*, *coil hair*, and *curvy hair*. Zarei (2002) reported that the most frequent problems of advanced-level Iranian learners of English with English collocations are in "noun+ verb" and the second most frequent errors are in "adjective+ noun". The finding had similarity with that in the present study because L2 adjective+ noun collocation errors were the second most frequent errors NEMs in KU made.

Secondly, the results revealed that one of the weaknesses of NEMs in KU to use collocation is the use of L4 quantifier+ noun pattern of collocations. As mentioned in Chapter 2, Farghal and Obiedat (1995) analyzed mis-collocations in students' collocation tests and declared that collocations especially those which are unpredictable from learners' first language, are difficult for them to master. Quantifiers in Chinese are different from those in English. For example, to describe

wind, applause, or laughter in Chinese “*yi zhen*” would be used but in English there are several quantifier candidates to express the same meaning, such as *a burst of*, *a ripple of*, *a gale of*, *around of*, *a gust of* and *a crash of*. Thus, to choose a proper and correct quantifier to go with a noun is difficult for NEMs in KU.

Finally, on the one hand, participants made no errors of L3 Noun+ verb collocations. This might be related to students’ avoidance of using this pattern in their composition. This type of lexical collocation consists of a noun and verb, the verb names an action characteristic of the person or thing designated by the noun (Benson et al, 1997). Examples are *bushes rustle*, *button bleep* and *mud squelch*. Words such as *rustle*, *bleep*, *squelch* are onomatopoeia, in other words they are words taken from sound. In Chinese, these words are usually used in spoken language. On the other hand, the participants made relatively fewer L5 Adverb+ adjective collocation errors than they made in L6 Verb+ adverb collocation, which pointed out that NEMs in KU had more trouble in choosing an adverb to go with a verb than choosing an adverb to go with an adjective. Such errors include *wholly acquainted*, *act simply*, and *celebrate gladly*. As stated before, NEMs in KU have problems to identify subtle differences between words especially verbs, probably based on their problems with verbs and then choose correct adjectives to go with verbs is a big challenge for them.

As mentioned in the previous section, twenty students who made the most frequent errors took part in a semi-structured interview. Based on the findings of the interview, the following reasons could be used to discuss why NEMs in KU made

unacceptable collocations in their writing. First, students translated Chinese into English literally according to the Chinese linguistic conventions. e.g. *They make **giant progress** <L2> (great progress) in English study.* Bahns (1993) investigated Polish and German EFL learner's performance in English collocations. He concluded that the majority of collocation errors can be traced to L1 influence. Similar conclusion was made by Farhal and Obiedat (1995), Al-Zahrani (1998), Liu (1999), Meng and Li (2005), and Fan (2009).

Second, some verbs such as *take, make, do* have a very wide applicability but have little meaning on their own, due to the lack of collocation knowledge some students have thought these verbs could replace one another. e.g. *He went back home to make a shower (take a shower).* Similar finding was stated in Liu (1999) that EFL learners may think that words such as *make, do, and take* are de-lexicalized verbs so they can replace one another freely. For example: *make some measures* instead of *take measures.*

Third, owing to lack of collocation knowledge, some students consider that verb with *-ing* and *-ed* endings function as adjectives. e.g. *informed article (informative article)*, and some learners use synonyms to replace node words or collocates. e.g. *They make **giant progress** <L2> (great progress) in English study.* Due to the lack of collocational knowledge, learners may ignore rules and restrictions and use incorrect vocabulary items or structures, such as use of a word coinage or use of a synonym for a lexical item in a collocation. It is seen as a "straightforward application of the open

choice principle” (Farghal & Obiedat, 1995, p. 323).

Finally, students were used to learning words in isolation, so they only understood the basic meanings of a word but did not know what words it would go with. They had learnt a certain number of words but still it was hard for them to make their sound natural. e.g. *She is a purely beautiful <L5> (flawlessly/ unspeakably beautiful) girl and studies hard.* Howarth (1998) reported that EFL learners made a lower percentage of conventional collocations but a higher percentage of deviant combinations. e.g. *perform a project, pay effort, and reach finding.* Other researchers such as Bahns and Eldaw (1993) and Farghal and Obiedat (1995) also found that L2 learners had a big gap between their receptive and productive knowledge of collocations. Meng and Li (2005), and Fan (2009) also stated that learners’ ability to use collocation is hampered by the deficiency of L2 lexis.

4.3 Effects of collocation instruction

An intact class of 40 students, who failed in the CET Band-6 in December, 2010 participated in the quasi-experiment, which was conducted during regular class time in an 8-week period. Among them 20 were male and 20 female, with an average age of 20.7 years old. They were from 5 different majors, and had been learning English for more than 10 years.

The first week, pre-test and pre-writing were given to the participants. The following six weeks were used for learning the six types of lexical collocations via using COCA (one class period for one week). In the last week, posttest, post-writing,

questionnaire, and interview were administered to the students (8 weeks in all). The quasi-experiment was carried out as scheduled and everything went on smoothly except that there was one participant who dropped out on the 6th week of instructional treatment. Therefore, there were a total of 39 participants who were third year NEMs in KU.

Aiming to use COCA to raise students' collocation awareness, the researcher selected several different content words in different parts of speech from students' text book (New Horizon College English Book 3) and their collocations were used as examples to introduce different types of lexical collocation to each class. Also, some passages and sentences were selected to design exercises for using COCA to raise students' lexical collocation awareness. Every class followed the same teaching steps and time allotment. Students were required to do fill-in-blank, matching, collocation collection, and error correction exercises designed according to Benson's lexical collocation types (one type for one class). At the end of each class, the researcher emphasized that student should do collocation collection exercise after class because the number of collocations is innumerable, and to some extent, the process of learning collocation is a process of collecting collocations. Lesson plan is shown in Appendix A.

4.3.1 Participants' Performance in Pre- and Post writings

For the writing tasks, on week 1 and week 8 of the quasi-experiment the participants were required to complete their writing within the time limit of one class

session (50 minutes). As a consequence, the data for the pre- and post writings were collected from 39 participants. All the data had been analyzed carefully by the same two raters following the error analysis procedure, which was presented in Chapters 1 and 2.

The number of lexical collocation errors that occurred in each students' composition was coded in SPSS 17.0 to analyze whether there is a significant difference between students' performance in pre-test and posttest.

Table 4.3 Comparison of Collocation Errors between Pre- and Post writing

Group	Scores	Mean	SD	N	Std. Error Mean	t	Mean Margin of Improvement	df	Sig.(2-tailed)
EG	Pre-writing	.62	.711	39	.114	1.302	+.23	38	.201
	Post-writing	.49	.601	39	.096				

Table 4.3 summarizes the results from Paired samples t-test. It shows that there was no significant difference between the pre-writing and post writing because $p = 0.201$ which is higher than the level of significance of 0.05. The mean for pre-writing is .69, while the mean for the post-writing is .49. A positive difference of .23 in the mean score indicates that there is a slight improvement in the post-writing.

Table 4.4 Correct and Incorrect Collocations in Pre- and Post writing

Type	Pre-writing		Post-writing	
	No. Correct collocation	No. Incorrect collocation	No. Correct collocation	No. Incorrect collocation
L1 Verb +noun	6	9	9	7
L2 Adjective +noun	7	5	7	5
L4 Quantifier +noun	2	3	4	2
L5 Adverb +adjective	3	5	3	2
L6 Verb +adverb	3	1	4	2
In all	21	23	27	18

As shown in Table 4.4, L1 and L2 are frequently used in learners' writing but the participants produced the same number of correct and incorrect collocations of L2 Adjective+ noun collocations in pre-writing and post-writing, totally 7 correct and 5 incorrect collocations. There is no L3 Noun + verb type of collocation found in students' pre- and post writing. For L1 Verb+ noun and L4 Quantifier + noun collocations, their performance improved because compared with pre-writing they produced more correct collocations but fewer incorrect collocations in post-writing. Compared with pre-writing, the participants performed better in using L5 Adverb + adjective patterns of collocations in post-writing, although they produced the same number of correct collocations in pre and post-writing, i.e. 3 collocations. The number of incorrect collocations reduced from 5 errors in pre-writing to 2 errors in post writing. As for L6 Verb + adverb type, in post-writing the participants produced such collocations more (4 correct and 2 incorrect) than in pre-writing (3 and 1), but they made one more error in post-writing. Therefore, it can be suggested that they

need to pay more attention to this kind of collocation in their future study.

To have some understanding of how serious these collocation errors were, the findings of lexical collocation error occurrences and percentage are listed in Table 4.5.

Table 4.5 Distribution of Lexical Collocation Error Domains

in Pre- and Post Writing

Tag	Pre-writing		Post-writing	
	No. of Occurrences	Percent (%)	No. of Occurrences	Percent (%)
L1	11	40.8	9	40.9
L2	7	25.9	7	31.8
L5	5	18.5	2	9.1
L4	3	11.1	2	9.1
L6	1	3.7	2	9.1
Total	27	100	22	100

Note: Error tags as shown in Table 3.2.

As shown in 4.5, in pre-writing the most frequent errors made were L1 errors. A total number of 11 tokens were found representing 9 L1 errors, which accounted for 40.8 % of the total errors. That is, 2 L1 errors occurred more than once. The most frequent errors in post-writing were L1 errors, too. A total number of 9 tokens were found representing 7 L1 errors, which accounted for 40.9%. That is, 1 L1 error was repeated in the learner corpus. In both pre- and post writings, the second most frequent type of errors was L2 errors. There were 5 errors realized by 7 tokens, which accounted for 25.9% and 31.8%, respectively. The L5 pattern of errors ranked the third frequent type in pre-writing, 5 errors which accounted for 18.5%. The L6 pattern of errors was the least frequent errors in the learner corpus, only 1 error in pre-writing

and 2 errors in post-writing, which accounted for 3.7 % and 9.1 %, respectively.

4.3.2 Participants' Performance on the Pre- and Posttest

Thirty-nine participants received collocation tests. Posttest was administered and the scores were compared with those of the pretest to determine the effects of the collocation instructional treatment. Both tests were scored according to the marking criteria presented in 3.2.2.2 of Chapter 3. Compared with the pretest, there were 4 students who even performed worse in the posttest, but the overall results were positive because the majority of students' performance improved in posttest.

Table 4.6 Comparison of Collocation Test Scores between Pre-test and Post test

Group	Scores	Mean	SD	N	Std. Error Mean	t	Mean Margin of Improvement	df	Sig.(2-tailed)
EG	pretest	11.69	4.098	39	.556	-2.303	+2.03	38	.027
	posttest	13.72	4.116	39	.559				

Table 4.6 summarizes the results from Paired samples t-test. It shows that there was a significant difference between the pre-test scores and posttest scores because $p = 0.027$ which is less than the level of significance of 0.05. The mean for pre-test score was 11.69, while the mean for the posttest was 13.72. A positive difference of 2.03 in the mean score indicates that there is an improvement in the posttest. It suggests that students made a significant improvement on their collocation performance after the 6-week instructional treatment.

4.3.3 Participants' Attitudes as Revealed by the Questionnaire

In response to the second research question: “What are the effects of using COCA to raise students’ lexical collocation awareness?”, the analysis of students’ questionnaire was considered as one of the important type of evidence.

In order to check the validity of all the questions in students’ questionnaire, 3 experts were invited to validate and check the language use for each item. As mentioned in Chapter 3, the content validity value of current questionnaire was 0.867 and the reliability from the pilot study was 0.83, which indicated the questionnaire was acceptable. After the instructional treatment, the students were required to answer the questionnaire. A 5-point Likert-scale questionnaire that ranged from “strongly disagree” to “strongly agree” was employed. Students’ responses to the questionnaires were coded and input into SPSS 17.0 for statistical analysis. The responses were described in Tables 4.7.

Table 4.7 Student’s Responses on the Likert-scale Questionnaires (N=39)

Item	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1 I think COCA is useful in searching for English collocations. (我认为 COCA 对查询英语词语搭配很有用)	2.6%	2.6%	17.9%	43.6%	33.3%
2 I think I will use COCA as an assistant tool to learn English collocations in the future. (我想将来我将使用 COCA 作为辅助工具来学习词语搭配)	2.5%	7.7%	20.8%	48.7%	17.7%
3 I think COCA is useful in helping me identify collocation errors in English. (我认为 COCA 能帮助我识别英语词语搭配的错误)	2.6%	2.5%	28.2%	43.6%	20.5%
4 I enjoy using COCA as support for learning English collocations. (我很喜欢用 COCA 来学习英语词语搭配)	2.5%	7.7%	41.0%	33.3%	12.8%
5 I think COCA is user-friendly. (我认为 COCA 很便于操作)	2.5%	25.4%	30.8%	31.0%	7.7%
6 I think I will recommend using COCA as support for learning English collocations to other students. (我想我将推荐 COCA 给其他同学学习英语词语搭配)	2.5%	2.6%	23.1%	38.5%	30.8%

Table 4.7 shows the responses from students' questionnaires in Part 3. On one hand, the majority of the students think that COCA is useful in identifying collocation errors and searching for English collocations, and they will not only use COCA in the future but also will recommend COCA to other students to learn English. For Item 1, the percentage of students who agreed that *I think COCA is useful in searching for English collocations* is 76.9% (43.6+33.3). For Item 6, the percentage of students who agreed that *I think I will recommend using COCA as support for learning English collocations to other students* is 69.3% (38.5+30.8). For Item 2, the percentage of students who agreed that *I think I will use COCA as an assistant tool to learn English collocations in the future* is 65.4% (47.7+17.7). For item 3, the percentage of students who agreed that *I think COCA is useful in helping me identify collocation errors in English* is 64.1% (43.6+20.5). On the other hand, as for Items 4 and 5, *I enjoy using COCA as support for learning English collocations* and *I think COCA is user-friendly*, more than 30% students were undecided. However, a negative aspect was mentioned and should be considered carefully. Some students considered that COCA is not user-friendly, and the percentage was 27.9% (25.4+2.5).

In addition, in the part of open-ended questions from students' questionnaire, the students reported that they had been learning English for more than ten years. A majority of students reported that they had received teachers' in-class support in using online resources, such as, Youdao online translator, Baidu Video, New Horizon

online learning course, and New Orient online English courses. Some students recommended their ways to learn collocation to other students, such as translating sentences, using notebook, and Association Memorizing Methods.

The findings suggest that after using COCA to raise students' collocation awareness for six weeks, the students performed slightly better in terms of the production of correct collocations in their writing and their performance in collocation tests. Furthermore, according to their response to the questionnaires and interview, their attitudes towards using COCA to learn collocations were positive. Therefore, COCA can be recommended as an assistant tool of teaching and learning vocabulary and collocations in the classroom.

4.4 Discussion

A number of factors could have contributed to the findings. First of all, there was significant improvement in posttest. Students who took part in the instruction treatment were trained to learn collocations by utilizing COCA in the class. As an assignment, after class they collected collocations and made notes about using COCA. Through the 6 weeks, the students mastered a certain number of collocations because the researcher did not teach the collocations in the tests. They performed better in the posttest. The finding accorded with Hsu's (2010) study in which treatment of collocation instruction promoted the students' performance on collocation tests. Hsu (2010) stated that direct collocation instruction could be worthwhile to be explored as

a teaching option although a follow up instructional study needs to be carried out to further support the findings. There was no significant improvement in post-writing. The present study was conducted for a period of only 8 weeks. Probably, this was because the time period was too short to allow such improvement to happen.

Secondly, perhaps the researcher was too “ambitious” to look for positive effects on two “demanding” writing tasks which tested productive skills. Probably, notable positive effects would require a much longer period of treatment. NEMs in KU reported in the interview, due to the lack of vocabulary and knowledge about which word could go with which word, it is hard for them to produce correct collocations. In order to improve students’ collocatin production, English teachers in KU should input more knowledge of collocation in classes. Zarei’s (2002) suggested that knowledge of collocation was largely ignored in case of Iranian learners of English and recommended that language teachers should introduce lexical items with their most frequent collocates.

Thirdly, Students’ attitudes towards using COCA to learning lexical collocation were positive. Research has shown that learning through a corpus to explore authentic language and obtain samples of texts can support the development of learners’ learning. Among various online resources, computerized corpora are popularly viewed to be able to facilitate inductive data-driven language learning in ways that have been difficult or impossible in the past (Johns, 1994; Leech, 1997; Kennedy & Miceli, 2001; Krishnamurthy, 2001).

Finally, it is difficult for EFL students to master numerous collocations. The

majority of the NEMs in KU reported that lexical collocation was hard to learn. Gitsaki (1999) also raised an interesting point that lexical collocations are more difficult to acquire compared to its counterpart, which is grammatical collocations. In addition, the majority of the participants reported that they thought COCA was useful in helping them identify lexical collocation errors in English. In order to know whether words could have been taught in collocations, Hodne (2009) conducted a study involving finding collocates of certain words and also checking frequency and strength of collocations. The findings indicate that COCA turned out to be useful to access this kind of information. As noted in the introductory text on the website, COCA allows you to search for exact words or phrases, lemmas, part of speech, or any combinations of these (Davies, 2008).

4.5 Summary of Chapter 4

In this Chapter, the results of error analysis, the collocation pretest and posttest, students' questionnaires, and pre-writing and post-writing were presented. The quantitative data were analyzed by paired samples t-test, and the distribution percentage was accounted. As to the qualitative data, students' collocation errors in learner corpus and responses of students' questionnaires were illustrated, respectively. From the analysis, two research questions have already been answered. The next Chapter will discuss the conclusion and implications of the present study in details.

CHAPTER 5

CONCLUSION AND IMPLICATIONS

In this chapter, the research findings are summarized and the conclusion is drawn according to the results of the study. The pedagogical implications to collocation learning and teaching are also presented. Finally, the limitations and suggestions for further research are described in details.

5.1 Conclusion of the Study

The purposes of the current study are to identify the lexical collocation errors made by NEMs in KU in their writing and to investigate the effects of utilizing COCA on raising learners' awareness concerning lexical collocations. The current study has tried to answer two research questions. First, what are the typical lexical collocation errors found in the writing of non-English major EFL learners? Second, what are the effects of using COCA to raise students' lexical collocation awareness?

As mentioned in Chapter 3, the participants were 150 NEMs in KU. They took part in CET Band-6 which was held in December, 2010, and were required to take two writing tasks. Following Benson et al's (1997) lexical collocation types, Corder's (1981) error analysis stage and Chow's (2006) error analysis procedure, error analysis was conducted manually by two raters. Typical collocation errors were detected,

identified, and classified by using COCA as a reference corpus. The results of error analysis were used to answer Research Question 1.

In general, the most frequent errors NEMs in KU made in their writings were lexical collocation errors with verbs as nodes, i.e. L1 Verb+ noun. The second most frequent errors were collocations with adjectives as nodes, i.e. L2 Adjective+ noun. And misuses of quantifiers were also found to be common in the corpus because the learners only produced 6 L4 Quantifier+ noun pattern of collocations, of which 4 were errors. According to the results of error analysis, the students made more errors in L6 Verb+ adverb pattern than in L5 Adverb+ adjective pattern. This indicates students had more trouble in choosing an adverb to go with a verb than choosing an adverb to go with an adjective.

Thirty-nine students who had passed CET Band-4 and failed in CET Band-6 formed an intact class who received the collocation instructional treatment, which was given one session (50 minutes) per week for six weeks. The focus of the instructional treatment was on using COCA to raise students' collocation awareness. The data were collected to answer Research Question 2 by administering the following to the participants: a pre- and post writing task, a pre- and post lexical collocation test, and a questionnaire. The pre- and post writings were designed to examine students' use of lexical collocation before and after the treatment. The lexical collocation tests were designed to investigate students' receptive and productive collocation knowledge. A questionnaire was designed to collect information about students' attitudes towards

using COCA to learn collocation and difficulty levels of different types of lexical collocations.

The results of error analysis from the pre- and post essays indicated that participants' use of lexical collocations improved in post-writing because they produced more correct collocations and fewer collocation errors in post-writing. The scores of pre and post collocation tests were compared and the findings showed that students achieved a significant improvement on their collocation performance after the 6-week treatment. Participants' answers to questionnaire were analyzed quantitatively and qualitatively, participants reported that L2 Adjective+ noun, L1 Verb+ noun, and L4 Quantifier+ noun are difficult to learn. This is in agreement with their performance in writing tasks and collocation tests. The findings of questionnaire suggested the majority of the students think that COCA is useful in identifying collocation errors and searching for English collocations, and they not only will use COCA in the future but also will recommend COCA to other students to learn English.

5.2 Pedagogical Implications

The present study has demonstrated that the participants experienced difficulties in using lexical collocation, resulting from inadequate collocation knowledge, and COCA could be used as an assistant tool to raise students' lexical collocation awareness. According to Nesselhauf (2003), collocations do deserve a place in language teaching. Based on the findings of this study, several pedagogical implications can be offered.

Firstly, the concept of learning vocabulary by the means of collocation should be exploited both in class and after class for the benefits of the students. The participants in the current study have realized the importance of collocation, but students who have never been trained before do not know what collocations are. Learning individual words and their meanings is not enough to achieve great fluency in a second language. Students who are used to learning vocabulary in isolation are hard to achieve native-like performance. Biskup (1992) pointed out that *noticing* and *becoming aware of* collocations might be the first steps leading to internalizing and learning collocations. EFL learners can effectively learn English by focusing attention on chunks or collocations instead of single words (Sinclair, 1991; Lewis, 1993). Therefore, vocabulary should be learnt by means of collocations both in class and after class, so that learners will notice “how words co-occur together”.

Secondly, corpora can be introduced to students to improve their ability to use collocations. As indicated in the present study, the participants who were trained to use COCA to learn collocations made progress in collocation use. Studies have shown that corpora play a crucial role in language learning (Kennedy & Miceli, 2001; Krishnamurthy, 2001). Besides COCA, there are other computerized tools providing analytical and effective methods to benefit EFL learners' English proficiency, such as, British National Corpus (BNC), English Taiwan Corpus (English TLC), Tango (a bilingual collocational concordancer), Collins COBUILD Concordance Samplers, and Collocations Sampler. These resources can provide authentic materials of target

language for EFL learners to consult and correct their errors and make their spoken and written English natural.

Thirdly, Teachers can expect students to produce more collocations by inputting more collocations in classes. Techniques, activities, and exercises can be employed to teach collocations. In Lewis' opinion (2002), meaning and message are primary but activities and exercises which can help learners to be aware of the linguistic aspect can aid in acquisition of the language.

Finally, in order to know students' weaknesses of using collocations, writing tasks and collocation tests can be employed. Once the collocation weaknesses have been identified, teachers can focus on these weaknesses when teaching collocations. As stated in Chapter 4, NEMs in KU made so many collocation errors in L1 Verb+ noun and L2 Adjective+ noun types of collocations, and it is hard for them to choose an appropriate adverb to go with a verb. Therefore, English teachers need to focus first on their weaknesses and provide more useful information for them to learn these types of collocations.

5.3 Limitations of the Study

In this study, theoretical triangulation involved error analysis and collocation instruction design. Methodological triangulation involved using quantitative and qualitative methods to collect data. The triangulated data collection included error analysis, pre- and posttest, pre- and post-writing, questionnaire. The triangulation

method in the study contributed to a better understanding of the investigation of NEMs' weaknesses of using collocation in KU and the effects of using COCA to raise EFL learners' collocation awareness, and enabled the researcher to verify the research findings. However, although the present study yielded some perspectives about using COCA to raise EFL learners' collocation awareness and shed some light on English language teaching and learning in Chinese colleges and universities, some limitations should be addressed.

Firstly, this study was conducted with a total number of 150 participants for the RQ 1. The size of the sample is small, and therefore the findings cannot be generalized to all NEMs in QMDAP in Guizhou Province. In addition, the generalizability of the findings of instructional treatment was also limited as the number of participants for the quasi-experiment is small too (only one class with a total number of 39 participants).

Another limitation is that collocations actually consist of lexical and grammatical ones. However, this study only focused on the use of lexical collocations. Consequently, the results obtained from the lexical collocation tests and instruction treatment may not actually reflect participants' complete knowledge of collocations. It is also important to note that the treatment only lasted for six weeks. If the treatment had been carried out over a longer period of time, the findings may have revealed a highly significant relationship between using COCA to learn collocations and students' ability to produce collocations.

Finally, the study only focused on how to use COCA to raise students' collocation awareness. The current study only analyzed collocations in the learner corpus but did not look into the quality of students' essays. As stated in Chapter 2, collocation is one of the obstacles for students to improve their writing skills. If the researcher probes deeply, the findings may reveal the relationship between collocation use and writing skills.

5.4 Recommendations for Further Studies

As mentioned in the literature review, research has shown that EFL learners face problems with collocations, resulting in producing erroneous collocations. The findings of this study are not conclusive. Thus, the need for more research on lexical collocations is urgent. Based on the limitations discussed in section 5.3 above, the researcher offers some recommendations for further studies in teaching and learning collocations in EFL contexts.

First of all, large scale studies covering more participants and a longer period of time are needed so as to make the findings more generalizable. Future researchers could consider a longitudinal study and investigate the development of collocation knowledge among EFL learners, similar to the one conducted by Gitsaki (1999).

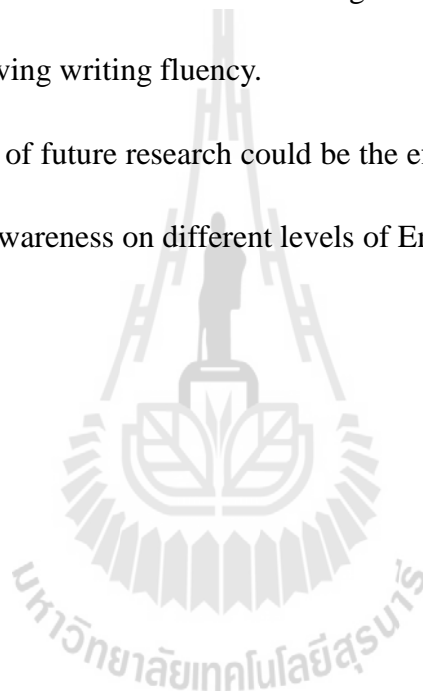
Second, future studies can be wider in scope to include all the possible lexical collocations, such as Noun + noun collocations, so that the data could be more representative of the participants' knowledge of collocations. Another area of

improvement over the present study could be to investigate the relationship between students' English learning experiences and their collocation knowledge.

Third, researchers in the future may add another dimension to the study, to investigate whether the collocation knowledge is sustained in a delayed post-essay writing or posttest.

Fourth, future research should consider grammatical collocations and its effectiveness in improving writing fluency.

Finally, the focus of future research could be the effects of using COCA to raise student's collocation awareness on different levels of English proficiency.





REFERENCES

- Abadi, S. H. (2003). A study of the learning of English lexical and grammatical collocations by Iranian EFL learners multiple-choice test. *Journal of faculty of letters and humanities (Tabriz)*, 46(187), 45-59.
- Al-Zahrani, M. S. (1998). *Knowledge of English lexical collocations among male Saudi college students majoring in English at a Saudi university*. Doctoral dissertation. Indiana University of Pennsylvania, Pennsylvania.
- Al-Sibai, D. (2009). *Testing Saudi English collocational competence and its development*. King Saud University. Retrieved from:
<http://docs.ksu.edu.sa/PDF/Articles06/Article060208.pdf>
- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELTJ*, 47(1), 56-63.
- Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21, 101-111.
- Baker, M. (1992). *In other words*. London: Routledge.
- Becker, J. (1995). The phrasal lexicon. In B. Nash-Webber & R. Shark (Eds). *Theoretical Issues in Natural Language Processing* (pp. 70-73). Cambridge: Bolt, Beranek.
- Benson, M., Enson, E., & Ilson, R. (1997). *The BBI Combinatory Dictionary of English (Rev. ed.)*. Amsterdam: John Benjamins Publishing Company.

- Benson, M. (1989). The structure of the collocational dictionary. *International Journal of Lexicography*, 2, 1-14.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In Sinclair, J. M. (Ed.), *How to use corpora in language teaching*, 15-36. Amsterdam: John Benjamins.
- Biskup, D. (1992). L1 influence on learner's rendering of English collocations: A polish/ German empirical study. In Arnaud, P. and Bejiont, H. (Eds.), *Vocabulary and applied linguistics*, 85-101. London: Macmillan.
- Boonyasaquan, S. (2006). An analysis of collocational violations in translation. *Journal of Humanities* (pp. 79-91). Bangkok: Faculty of Humanities, Srinakharinwirot University.
- Bonk, W. (2000). *Testing ESL learners' knowledge of collocations*. U.S.: Illinois. (ERIC Document Reproduction Service No. ED442309)
- Booncherd, P. (1974). *Evaluation by References Criteria, Concepts and Methods*. Faculty of Educational Study, Srinakrindravivot University (Ed.) *Fundamental Education Division* (pp: 89-299). Bangkok: ISBN: 974-275-874-3.
- Carter, R. (1998). *Vocabulary: Applied Linguistic Perspective*. London: Routledge.
- Carter, R., & McCarthy, T. (1988). *Vocabulary and language teaching*. New York: Longman.
- Chambers, A., & O'Sullivan, I. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL*, 16(1), P 158-172.

- Chen, W. X. (2002). Collocation errors in the writings of Chinese learners of English. *Journal of PLA University of Foreign Language*, 25(1), 60-62.
- Church, K. & Patrick, H. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1), 22-29.
- Chow, P. K. (2006). *Tense and aspect in interlanguage: Error analysis in the English of Cantonese-speaking secondary school students*. Master degree thesis. University of Hong Kong.
- Clear, J. (1993). From Firth principles – Computational tools for the study of collocation. In Baker, M., Francis, G., & Tognini-Bonelli, E. (Eds). *Text and technology: In honour of John Sinclair*, 271-292. Amsterdam: John Benjamins.
- Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford: OUP.
- Cowie, A. P. (1994). Phraseology. In *The encyclopedia of language and linguistics*, Asher, R.E., & Simpson, J. (Eds), 3168-71. Oxford: Pergamon Press.
- Crowther, J. (2001). *Oxford Collocations, Dictionary for Students of English*. Oxford: Oxford University Press.
- Davies, M. (2008). Corpus of Contemporary American English (COCA). Retrieved from: <http://www.americancorpus.org>.
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). London: Sage Publications, Inc.
- Ellis, R. (1994). *Understanding Second Language Acquisition*. Shanghai Foreign Language: Education Press.

- Fan, M. (2009). An exploratory study of collocational use by ESL students-A task based approach. *System*, 37, 110-123.
- Farghal, M., & Obiedat, H. (1995). Collocation: A Negative Variable in EFL Writing. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 33(4), 315-331.
- Gitsaki, C. (1999). *Second Language Lexical Acquisition: A study of the development of collocational knowledge*. International Scholars Publications (San Francisco, CA).
- Goulden, R., Nation, P. & Read, J. (1990). How Large Can a Receptive Vocabulary Be? *Applied Linguistics*, 11(4), 341-363.
- Halliday, M. A. K. & Hason, R. (2001). *Cohesion in English*. Beijing: Foreign Language Teaching and Researching Press.
- Hill, J. (2000). *Revising priorities: From grammatical failure to collocational success*. In Lewis, M. (Ed.), *Teaching Collocation, Further Development in the Lexical Approach* (pp. 47-69). London: Heinle, Cengage Learning.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 4, 24-25.
- Hodne, L. C. S. (2009). *Collocations and Teaching*. Master dissertation, University of Bergen.
- Hsu, J. (2010). The effects of collocation instruction on the reading comprehension and vocabulary learning of Taiwanese college English majors. *The Asian EFL Journal Quarterly*, 12(1), 47-87.

- James, C & Thomas, H. (1997). *Second Language Vocabulary Acquisition: A rationale for pedagogy*. New York : Cambridge University Press.
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. London: Longman.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of Data-driven Learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 293-313). Cambridge: Cambridge University Press.
- Keshavarz, M. H. & Salimi, H. (2007). Collocational competence and cloze test performance: a study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17(1), 81-92.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, 5 (3), 77-90.
- Kjellmer, G. (1987). Aspects of English collocations. In Willem Meijs, editor, *Corpus Linguistics and Beyond*, pp. 133–140. Rodopi, Amsterdam.
- Kjellmer, G. (1990). Patterns of collocability. In Aarts, J., & Meijs, W. (Eds.), *Theory and practice in corpus linguistics*, 163-178. Amsterdam: Rodopi.
- Kjellmer, G. (1991). A mint of phrases. In Aijmer, K., & Altenberg, B. (Eds.), *English Corpus Linguistics*, 111-127. London: Longman.
- Krishnamurthy, R. (1997). *The Process of Compilation*. London: Language Teaching Publications.

- Krishnamurthy, R. (2001). Language corpora: How can teachers and students use these valuable new resources? *Selected Papers from the 10th International Symposium on English Teaching*, 59-65. Taipei: Crane.
- Lancaster, G. A., Dodd, S., and Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10 (2), 307.
- Leech, G. (1997). Teaching and language corpora: A convergence. In Wichmann, A., Flingelstone, S., McEnery, T., & Knowles, G. (eds.) *Teaching and Language Corpora*, pp. 1-23. London and New York: Longman.
- Lewis, M. (2000). *There is nothing as practical as a good theory*. In Lewis, M. (Ed.), *Teaching collocation*, 10-27. London: Heinle, Cengage Learning.
- Lewis, M. (2002). *Implementing the Lexical Approach: Putting Theory into Practice*. Hove UK: Language Teaching Publication.
- Lewis, M. (1993). *The Lexical Approach: The state of ELT and a way forward*. Hove UK: Language Teaching Publication.
- Lien, H-Y. (2003). *The effects of collocation instruction on the reading comprehension of Taiwanese college students*. Unpublished doctoral dissertation, Indiana University of Pennsylvania, Pennsylvania.
- Liu, C. P. (1999). A study of Chinese Culture university freshmen's collocational competence: "Knowledge" as an example. *Hwa Kang Journal of English Language & Literature*, 5, 81-99.

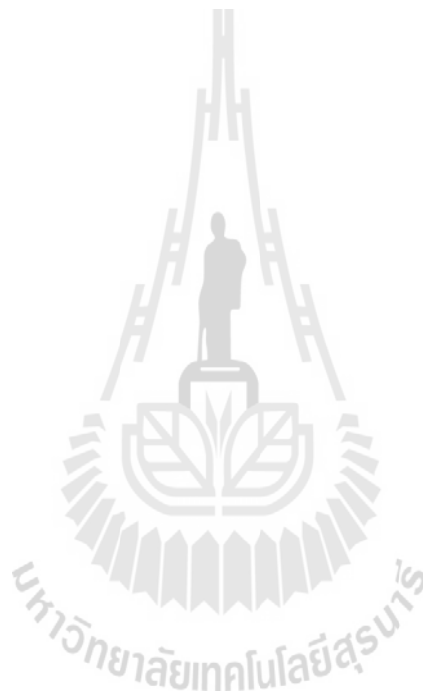
- Malligamas, P. & Pongpairroj, N. (2005). Thai learners' knowledge of English collocations. *HKBU Papers in Applied Language Studies*, 9, 1-28.
- Mackin, R. (1978). *On collocations: Words shall be known by the company they keep*. In P. Strevens (Ed.), *In Honour of A. S. Hornby*, 149-165. Oxford: Oxford University Press.
- McEnery, T., Xiao, R. & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. New York: Routledge.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Meng, J., & Li, L. (2005). A Tentative Analysis of Collocational Errors in Chinese EFL Learners' Composition. *Journal of Shangdong Electric Power Institute*, 3, 20-22.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House Publishers.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24, 223-242.
- Nunan, D. (2001). *Research methods in language learning*. Cambridge: Cambridge University Press.
- Oxford Collocations Dictionary* (2002). (7th ed.), Oxford: Oxford University Press.
- Richards, J. (2001). *Curriculum Development in Language Teaching*. Cambridge: Cambridge University Press.

- Robson, C. (2002). *Real World Research: A resource for social scientists and practitioner---researchers*. Oxford: Black well Publish.
- Seidlhofer, B. (2000). Operationalizing intertextuality: Using learner corpora for learning. In McEnery, T. & Lou, B. (Eds.), *Rethinking language pedagogy from a corpus perspective*, 207-233. Frankfurt: Peter Lang.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two versions of vocabulary levels test. *Language Testing*, 18, 55-88.
- Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Smadja, F., & McKeown, K. (1991). Using collocations for language generation. *Computational Intelligence*, 7, 229-239.
- The National College English Testing Committee (NCETC) (n.d.). *College English Test Band-4 and Band-6*. Retrieved from: <http://www.en.cet.edu.cn/>
- Thomas, R.M. (2003). *Blending qualitative and quantitative research approach to language learning*. Englewood cliffs, HJ: Prentice Hall.
- Wang, H. Y. (2010). An analysis of obstacles of undergraduates' listening and speaking in ethnic communities. *Journal of Kaili University*, 28 (2), 79-81.
- Webb, S. & Kagimoto, E. (2010). Learning Collocations: Do the Number of Collocates, Position of the Node Word, and Synonymy Affect Learning? *Applied Linguistics*, 1–19. Oxford University Press.

- Wu, Z. (2003). Developing students writing strategies: an investigation of language errors from students' compositions. *Journal of Southeast Guizhou National Teacher's College*, 21 (1), 114-115.
- Wray, A. (2005). *Formulaic Language and Lexicon*. Cambridge: Cambridge University Press.
- Wei, Y (1999). Teaching Collocations for Productive Vocabulary Development. *Annual Meeting of the Teachers of English to Speakers of Other Languages*, 9-13. New York.
- Xie, F., & Jiang X. M. (2007). Error analysis and the EFL classroom teaching. *US-China Education review*, 4(9), 10-14.
- Yang, H., & Weir, C. (1998). *Validation study of the national college English test*. Shanghai: Shanghai Foreign Language Education Press.
- Yi, X. (2004). *Applied Linguistics: Language learning and teaching*. Changchun: Jilin University Press.
- Ying, Y. & Hendricks, A. (2004). Collocation awareness in the writing process. *Reflections of English Language Teaching*, 3, 51-78.
- Zarei, A. (2002). Patterns of Iranian advanced learners problems with English collocations. Unpublished doctoral dissertation, University of Isfahan, Iran.
- Zheng, H. P. (2000). On today's English reading problems and improvement measures. *Journal of Southeast Guizhou National Teacher's College*, 18(1), 84-85.

Zorana, V. (2008). Developing Collocational Competence of Second Language Learners. *The East Asian Learner*, 4(1), 46-50.

Zughoul R., & Hussein, A. (2003). Collocational Strategies of Arab Learners of English: A Study in Lexical Semantics. Department of English, Yarmouk University, Irbid, Jordan, *BABEL*, 49, 1. (ERIC Document Reproduction Service No. ED479746)





APPENDIX A

Lesson Plan for Teaching Lexical Collocation

Objectives:

At the end of the lesson, students will be able to:

1. query lexical collocations in COCA;
2. find collocates to given nodes;
3. find correct lexical collocations from concordance lines for correcting lexical collocation errors.

Teaching & learning focus:

1. What is lexical collocation?
2. Benson's lexical collocation types.
3. Each week for each type (Week 2: Verb+ noun; Week 3: Adjective+ noun; Week 4: Noun+ verb; Week 5: Quantifier+ noun; Week 6: Adverb+ adjective; Week 7 Verb+ adverb).

Materials:

1. Video 1: How to use COCA? Corpus of Contemporary American English Guide/Tutorial: <http://www.youtube.com/watch?v=jNRBjUL2s90> or <http://vimeo.com/7218683>
2. Video 2: What is collocation?
Introduction to Collocations http://www.youtube.com/watch?v=dss_hlqmMGM
3. Online corpus COCA: <http://corpus.byu.edu/coca/>
4. New Horizon College English Book 3 (Reading and Writing)

Duration: 50 minutes

Teaching content and arrangement:

Stages	Approx. time	Activities	Purposes
Warming-up	3 minutes	Step 1: Greeting and leading-in	To help students get mentally prepared for learning
Watching Video 1	8 minutes	Step 2: Students watching Corpus of Contemporary American English Guide/Tutorial video	To help students get a clear picture how to use COCA
Watching Video 2	4 minutes	Step 3: Students watching video: What is collocation?	To help students to know what is collocation and how to learn collocation

Stages	Approx. time	Activities	Purposes
Lexical collocation and its types	6 minutes	Step 4: Teacher explaining what lexical collocation is with examples (select words in New Horizon College English Book 3) Step 5: Teacher showing Benson et al's lexical collocation types and explaining these types to students	Steps 4 and 5: to help students get to know the concept of lexical collocation and what are lexical collocation types
Activities	24 minutes	Step 6: Using COCA to do three collocation learning exercises: 1. Fill-in-blank exercise 2. Matching exercise 3. Lexical collocation error correction exercise Step 7: Students doing peer correction activity Step 8: Based on exercises 1 and 2, students use COCA to make notes for collecting lexical collocations	Step 6: to help students to learn how to query lexical collocations in COCA and find correct lexical collocations from concordance lines for correcting lexical collocation errors Steps 7 and 8: to help students to learn which word can go with which word
Reflection	5 minutes	Step 9: Wrap up Students and teacher discussing what they have learnt during class	To help consolidate students' collocation knowledge

Note: The paragraphs or sentences in Exercise 1 and 2 are selected from the textbook: *New Horizon College English Book 3 (Reading and Writing)*. Items in the error correction exercises are selected from collocation errors students made.

Week 2

Activities: Using COCA to learn L1 Verb+ noun collocation

Exercise 1: Fill-in-blank

A, filling in the blanks

In ___ a story, the easiest and clearest way is to ___ in details of what happens in order of time. This method is called chronological sequencing. Here is a topic for you to write a paragraph by ___ chronological orders: On weekend.

Key: telling describe following

B. Lexical collocation collection.

tell: story/truth/tales/jokes/pollsters/anecdote...

describe: details/phenomena/methodology/scenarios/ecstasy...

follow: orders/ rules/ path/lead/suit/directions/instructions/guidelines...

Exercise 2: Matching

A. send take avoid

Students who usually ___ making speaking mistakes in class become active in the online environment. Participants are more willing to ___ the chance of written participation than speaking, perhaps partly because they can rethink before

e-mails.

Key: avoid take sending

B. Lexical collocation collection

avoid: mistakes/conflict/pitfalls/confusion...

take: chance/advantage/care/action/steps...

send: e-mails/message/letter/signal...

Exercise 3: Lexical collocation error correction

Identify a collocational error in each sentence by underlining it out, and then correct it.

We can't fit the demand of finding a job.

Correction: to meet the demand

Many students open the lights even though they go out to play.

Correction: turn on the lights

Students learn knowledge from their teachers.

Correction: acquire/gain knowledge

Week 3

Activities: Using COCA to learn L2 Adjective+ noun collocation

Exercise 1: Fill-in-blank

A. filling in the blanks

The ability to tell the difference between facts and writer's opinions plays a/an role in reading. Facts are usually based on ____ evidence. Opinions are often based on facts, but they also involve the writer's ____ opinions of the facts, which may or may not match your opinions of them.

Key: crucial direct personal

B. Lexical collocation collection

crucial: role/ element/ success/ factor/ step...

direct: evidence/ contact/ instruction/effects...

Personal: opinion/ experience/ income/ style...

Exercise 2: Matching

A. main minor vast

Words in a sentence describe ____ details about the subject of the sentence. Although a paragraph may give ____ quantities of information, it usually offers one idea.

Key: minor vast main

B. Lexical collocation collection.

minor: details/leagues/changes/adjustments...

vast: quantities/amount/number/expense...

main: idea/street/road/reason/effect...

Exercise 3: Lexical collocation error correction

Identify a collocation error in each sentence by underlining it out, and then correct it.

In order to have an attracting looking they spend much money.

Correction: an attractive looking

Our country has grand population in cities.

Correction: large population

Few of them can speak smooth English.

Correction: fluent English

Week 4

Activities: Using COCA to learn L3 Noun+ verb collocation

Exercise 1: Fill-in-blank

A. *filling in the blanks*

There was a loud boom as the bomb _____.

The refrigerator _____ softly in the corner.

Thunder _____ across the sky.

Key: *exploded hummed rumbled*

B. *Lexical collocation collection*

bomb/car/challenger/shell...exploded

refrigerator/machine/conditioner/engine...hummed

thunder/train/stomach/truck...rumbled

Exercise 2: Matching

A. *tolled buzzed rustled*

Leaves _____ in the breeze.

The shaver _____.

The church bells _____.

Key: *rustled buzzed tolled*

B. *Lexical collocation collection*

leaves/ wind/ grass/branches...rustled

shaver/ flies/ mosquito/ bees/insects...buzzed

bells/clock/message...tolled

Exercise 3: Lexical collocation error correction

Identify a collocation error in each sentence by underlining it out, and then correct it.

Their glasses ringed.

Correction: glasses clinked

The loudspeaker squeaked.

Correction: loudspeaker crackled.

I could hear the coins banged in his pocket.

Correction: coins jingled

Week 5

Activities: Using COCA to learn L4 Quantifier+ noun collocation

Exercise 1: Fill-in-blank

A, *Filling in the blanks*

Would it be wrong for someone to steal _____ bread to feed his starving family? She is the youngest of _____ seven children.

_____ bees is flying among the flowers.

Key: a loaf of bread/ a brood of children/ a cluster of bees

B. *Lexical collocation collection*

a loaf of: bread/ rye/ wheat/ sourdough...

a brood of: children/ chicks/ puppies/ mongooses...

a cluster of: bees/ building/ brick/ galaxy...

Exercise 2: Matching

A. *a pack of dogs / a branch of knowledge/ a flock of sheep*

The pupil huddled together like _____sheep.

If you encounter _____dogs on the street, run!

I am running short on cash. I wonder if there's _____Agriculture Bank around.

Key: *a flock of sheep/ a pack of dogs/ a branch of knowledge*

B. *Lexical collocation collection.*

a flock of: sheep/ birds/ geese/ pigeons...

a pack of: dogs/ cigarettes/ wolves/ lies...

a branch of: bank/ government/ department/rive...

Exercise 3: Lexical collocation error correction

Identify a collocation error in each sentence by underlining it out, and then correct it.

With a little of humor, she told us her story.

Correction: a sense of / a dash of/ a bit of humor

You'll get a break of energy like you've never known before.

Correction: a burst of energy

I still have a line of hope I can earn money to pay the tuition fee by my own, but parents didn't think so.

Correction: a glimmer of hope

Week 6

Activities: Using COCA to learn L5 Adverb+ Adjective collocation

Exercise 1: Fill-in-blank exercise

A. *Try to use different adverbs to fill in the blanks*

On the ship he meets a/an _____beautiful girl named Isabella. One night Isabella can't sleep, so she goes for a walk on the ship. She meets John and they start talking. They talk all night and then fall in love with each other. John is a/an nice person. But Isabella is going to marry a/an_____rich man.

Key: *incredibly awfully extremely*

B. *Lexical collocation collection*

incredibly/strikingly/ hauntingly/ extraordinarily...beautiful

awfully/ perfectly/ really/ extremely...nice

extremely/ particularly/ especially/ fabulously...rich

Exercise 2: Matching

fully purely truly

Online learning is _____sequential program. It requires commitment on the students' part. Remember that instructors cannot see their students in an online

course. This means students must be _____ explicit with their comments and requests. If they experience difficulties which are _____ technical problems they must speak up, otherwise there is no way anyone can know something is wrong.

Key: truly absolutely purely

B. Lexical collocation collection

truly/ less/ strictly/ necessarily...sequential

quite/ spatially/ fully/ partially...explicit

purely/ highly/ merely/ fairly...technical

Exercise 3: Lexical collocation error correction

Identify a collocational error in each sentence by underlining it out, and then correct it.

She is a purely beautiful girl and studies hard.

Correction: flawlessly/ unspeakably beautiful

We are wholly acquainted after living together for a month.

Correction: thoroughly/closely/intimately acquainted

Their performance was greatly amusing on birthday party.

Correction: highly amusing

Week 7

Activities: Using COCA to learn L6 Verb+ adverb collocation

Exercise 1: Fill-in-blank

A. Filling in the blanks

Job opportunities for students are expanding _____ and more people of all ages are becoming aware of online learning that allows them to study at home. The forum for communication removes the visual barriers that hinder some students from expressing themselves and encourages students to take the program _____. In addition, students are given time to reflect on the information before replying. In this way, students can help to keep the online environment open _____.

Key: rapidly seriously wide

B. Lexical collocation collection

expand: rapidly/ forever/ outward/ dramatically...

take: seriously/ away/ kindly/ awhile...

open: wide/ daily/ indefinitely/ automatically...

Exercise 2: Matching

clearly directly strongly

Writers sometimes express all their thoughts _____. The main idea might be stated in a sentence in a paragraph. However, it is also possible for the main idea is implied in a passage by the author.

Key: directly clearly strongly

B. Lexical collocation collection

express: directly/publicly/ freely/ candidly...

stated: clearly/ explicitly/ flatly/above....

implied; strongly/ above/ accordingly/ effortlessly...

Exercise 3: Lexical collocation error correction

Identify a collocational error in each sentence by underlining it out, and then correct it.

My shoes fit with my clothes wholly.

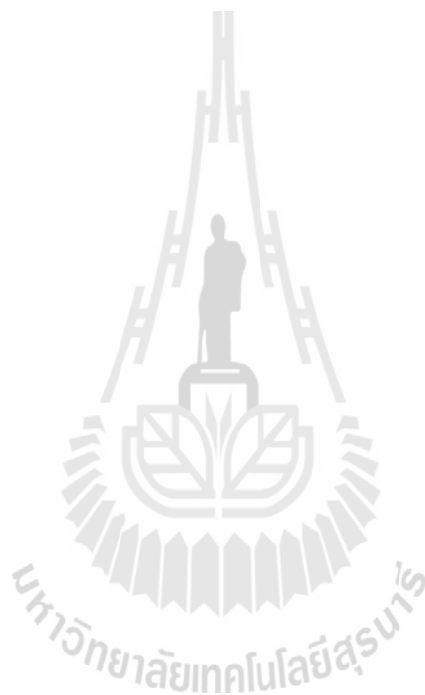
Correction: fit/perfectly/nicely/neatly/properly

They may act simply after drinking so much wine.

Correction: act rashly

We can celebrate our birthday gladly .

Correction: celebrate joyously



APPENDIX B

Collocation Test

_____ Institute of Kaili University Name _____

Date _____

Score _____

Part 1: Matching

Match the words in the box with the following sentences to make the context meaningful, you can use each word only once.

*take ecological fact overcome alarms set extremely heatedly ignite
cloud exchange boom intellectual humbly completely*

1. You should _____ the problem and not mention about it.
2. John ignores the _____ that men behave differently than women.
3. They _____ a debate over a variety of genetic birthrights and women's rights.
4. _____ property results from creative thinking and may include products, services processes, and ideas.
5. Central banks need the funds to intervene in the foreign _____ market.
6. The baby _____ swelled the population.
7. Every morning when _____ go off, he gets up to dress.
8. Jim fishes a/an _____ of keys from his pocket and tosses them to Shane.
9. All I can see is a/an _____ of dust far off in the distance.
10. It was _____ difficult to perform the act.
11. I am sorry I _____ forgot your birthday.

12. They both get out of their cars and argue _____.
13. They were waiting for her to explain why and to apologize _____.
14. Mama sent us to _____ a nap that afternoon after lunch.
15. The Foundation seeks to protect _____ balance for future generations.

Part 2 Error Correction

There is collocation error (an error caused by wrong choice of words) in the following sentences. Identify the wrong word in each sentence by underlining it out, and then correct it.

For example: She can do anything in order to succeed her goal. (achieve)

16. In 1915 Einstein made a trip to Gattingen to make some lectures.
17. I was content to treasure every moment of every minute just being with her.
18. Greggs has already taken measures to explain the problem.
19. He was a strong smoker and died of lung cancer at age 59.
20. On National Day the whole country is immersed in a festival situation.
21. The Gulf of Mexico oil drop kills fish and birds as well as pollutes the sea.
22. Most of us have seen the airplane descend at Hongqiao international airport.
23. Police had shown him a suit of glasses.
24. Flight refers to a mob of birds in flight.
25. Indeed, the lake near the factory is strongly polluted.
26. Rwanda is the most thickly populated country in Africa.
27. Somebody should wake him up. He's really snoring deeply.
28. Home prices are adding rapidly.
29. She always likes to talk jokes and get people laughing.
30. The two parties tried to seek general grounds while reserving differences.

APPENDIX C

QUESTIONNAIRE

问卷调查

This survey questionnaire is intended only for research purposes. It is designed to gather information about your opinions on the difficulty level of different types of collocation, your attitudes toward using COCA to learn collocation, and your English learning experience and recommendation. Please complete the following questions to reflect your attitudes as accurately as possible according to the specifications of each part. Your answers will be kept strictly confidential. Your help will be greatly appreciated.

本问卷主要是想了解各位学生对不同类型英语词语搭配难易度的观点,对使用美国当代英语语料库(COCA)学习词语搭配的看法,以及您的英语学习经历和对学习英语词语搭配的建议。此问卷仅作为学术研究资料之用,请放心作答,感谢您的合作!

Part 1: Your Background Information (基本信息)

Name (optional) 姓名(可随意填写): _____

Gender 性别: male 男 female 女

Major 专业: _____ Age 年龄: _____

Part 2: Your Attitudes Toward Using COCA to Learn Collocation. (您对使用COCA学习词语搭配的看法)

For this part, please read each statement carefully and check (✓) the response which best describes your opinions. The numbers 1 to 5 stand for the following responses:

请仔细阅读以下陈述并根据每个数字所代表的回答勾出您对使用COCA学习词语搭配的看法。

1 = Strongly Disagree 完全不同意 2 = Disagree 不同意 3 = Undecided 不确定 4 = Agree 同意 5 = Strongly Agree 完全同意

N O.	Your attitudes toward using COCA to learn collocation. (您对使用美国当代英语语料库学习词语搭配的看法)					
1	I think COCA is useful in searching for English collocations. (我认为 COCA 对查询英语词语搭配很有用)	1	2	3	4	5
2	I think I will use COCA as an assistant tool to learn English collocations in the future. (我想将来我将使用 COCA 作为辅助工具来学习词语搭配)	1	2	3	4	5
3	I think COCA is useful in helping me identify collocation errors in English. (我认为 COCA 能帮助我识别英语词语搭配的错误)	1	2	3	4	5
4	I enjoy using COCA as support for learning English collocations. (我很喜欢用 COCA 来学习英语词语搭配)	1	2	3	4	5
5	I think COCA is user-friendly. (我认为 COCA 很便于操作)	1	2	3	4	5
6	I think I will recommend using COCA as support for learning English collocations to other students. (我想我将推荐 COCA 给其他同学学习英语词语搭配)	1	2	3	4	5

Part 3: Your English learning experience and Recommendation. (您的英语学习经历和建议)

Please read each question carefully, give them brief answers or check (✓) the response to the following questions. 请仔细阅读每一个问题，然后作出简短回答或是勾出您的答案。

1. How long have you been learning English? 您学习英语有多久了?

5-9 Years (5-9年) 10-14 Years (10-14年) More than 15 Years (超过15年)

2. Have you ever received teachers' in-class support in using online resources? If Yes, what are they? 您的老师曾经在课堂上使用网上资源教学吗? 请勾出您的答案, 如果答案是“是”请写出使用哪些资源?

Yes是

No不是

Online

resources

3. Do you have your ways to learn collocation? If yes, please recommend them to

other students. 您有自己的词语搭配学习方法吗？如果有请把它推荐给其他同学。

Yes 有 No 没有

Recommendations (我的词语搭配学习方法) :

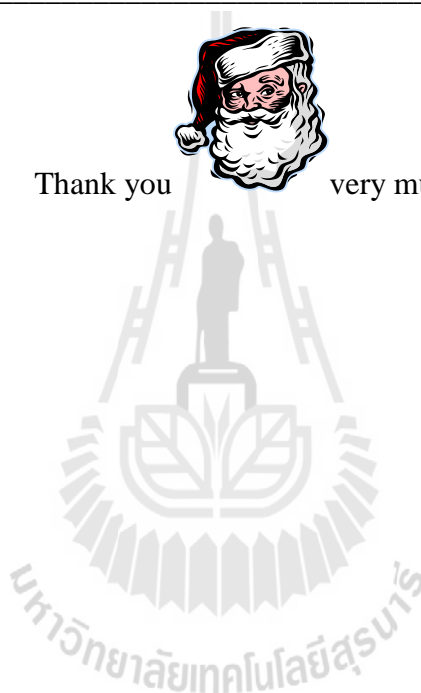
a. _____

b. _____

c. _____



Thank you very much!



APPENDIX D

Questions for Interview

Questions for students who made collocation errors (01-20):

- 1) What made you make the error “.....”?
你为什么犯“.....”这个搭配错误?
- 2) Why is it difficult for you to master such collocation?
为什么这种搭配难掌握?
- 3) In your opinion, how can you correct the error?
如何才能纠正这个错误呢?

Two more questions for students who come from the intact class and made collocation errors (01-06):

- 4) Do you think COCA is good for improving your English study? Why or why not?
你认为美国当代英语语料库能提高你的英语学习吗? 为什么?
- 5) Do you think that COCA should be used in English classes? Why or why not?
你认为美国当代英语语料库应该应用于英语课堂中吗? 为什么?

Interviewees:

Student	Errors (type)	QQ number
01	learn knowledge(L1)	164---4724@qq.com
02	coil hair(L2)	456---659@qq.com
03	a line of hope(L4)	950---146@qq.com
04	strongly polluted (L5)	678---442@qq.com
05	drive tiredly(L6)	412---923@qq.com
06	eat tea (L1)	112---2375@qq.com
07	make advantage(L1)	221---3195@qq.com
08	do a call(L1)	456---6783@qq.com
09	smooth English(L2)	865---567@qq.com
10	attracting looking(L2)	657---569@qq.com
11	insistent efforts (L2)	783---164@qq.com
12	a little of humor (L4)	653---801@qq.com
13	a bundle of flowers (L4)	341---114@qq.com
14	a cloud of wind (L4)	451---890@qq.com
15	definitely value (L5)	077---245@qq.com
16	purely beautiful (L5)	469---211@qq.com
17	totally acquainted (L5)	356---012@qq.com

18	manage reasonably (L6)	133---577@qq.com
19	arrange properly (L6)	346---923@qq.com
20	fit wholly (L6)	812---567@qq.com

APPENDIX E

Collocation Errors in the Learner Corpus and Corrections

Type	Collocation Errors and Corrections	Number (%)
L1Verb+ noun	<p>1. learn knowledge (acquire knowledge); 2. study knowledge (gain knowledge); 3. plant consciousness (raise consciousness); 4. eat tea (drink tea); 5. fit the demand (meet the demand); 6. add burden (compose burden); 7. open the light (turn on the light); 8. close the light (turn off the light); 9. make advantage (take advantage); 10. do plans (make plans); 11. treasure every moment (cherish every moment); 11. less waste (reduce waste); 12. provide chances (provide opportunity); 13. build consciousness (raise consciousness); 14. prove promise (fulfill promise); 15. make measures (take measures); 16. come to university (enter university); 17. order opportunity (offer opportunity); 18. get students achievement (improve students achievement); 19. display ability (demonstrate ability); 20. exhibit ability (demonstrate ability); 21. commit fire (commit arson); 22. do a call (make a call); 23. make the laundry (do laundry); 24. make our best (do our best); 25. do an offer (make an offer); 26. do negotiation (conduct negotiation); 27. make advantage (take advantage); 28. get a passion (follow a passion); 29. get career path (follow career path); 30. gave ccommodation (offer accommodation); 31. solve foreign affairs (conduct foreign affairs); 32. get interview (have interview); 33. open antibiotic (prescribe antibiotic); 34. raise anxiety (cause anxiety); 35. lose an appointment (cancel an appointment); 36. vote ballot (cast ballot); 37. use a bandage (apply a bandage); 38. do a perm (get a perm); 39. withdraw a permit (revoke a permit); 40. create a phrase (coin a phrase); 41. join in university (enter university); 42. perform fair (play fair); 43. break a plot (foil a plot);</p>	85 (50.9%)

	44. draw feathers (pluck feathers); 45. extract flowers (pluck flowers); 46. put plug (insert plug); 47. underline a point (emphasize a point); 48. manage	
Type	Collocation Errors and Corrections	Number (%)
	pollution (control pollution); 49. have popularity (gain popularity); 50. keep his popularity (maintain his popularity); 51. appreciate popularity (enjoy popularity); 52. produce possibility (raise possibility); 53. overcome possibility (exclude possibility); 54. use his poultice (apply his poultice); 55. oppose pressure (resist pressure); 56. awarded privilege (granted privilege); 57. made proclamation (issued proclamation); 58. display proficiency (demonstrate proficiency); 59. employed a program (implement a program); 60. mention a reference (contain a reference); 61. make redress (seek redress); 62. recruit volunteers (appeal volunteers); 63. create a record (establish a record); 64. overstep record (surpass record); 65. provide a recommendation (write a recommendation); 66. accept the actuality (accept the reality); 67. need a ransom (demand a ransom); 68. emit radiation (eject radiation); 69. indicate rage (express rage); 70. demonstrated prowess (displayed prowess); 71. arrive puberty (reach puberty); 72. give punishment (administer punishment); 73. make a rite (perform a rite); 74. equipped a room (furnished a room); 75. express a role (perform a role); 76. take sacrifices (make sacrifices); 77. perform sadness (express sadness); 78. furnish sanctuary (provide sanctuary); 79. arouse a scandal (cause a scandal); 80. disclose a scandal (uncover a scandal); 81. polish a knife (sharpen a knife); 82. do a seat (take a seat); 83. decide sentence (pronounce sentence); 84. implement ordinance (enforce ordinance); 85. indicate outrage (express outrage)	
L2 Adjective +noun	1. smooth English (fluent English); 2. smeary environment (polluted environment); 3. social sign (social phenomenon); 4. attracting looking (attractive looking); 5. healthing environment (healthy environment); 6. insistent efforts (ceaseless efforts); 7. environmental balance (ecological balance); 8. coil hair (curly hair); 9. giant progress (great progress); 10. informed article (informative article); 11.	46 (27.5%)

	communicative understanding(mutual understanding); 12. uneven state (uneven terrain); 13. apace development (rapid development); 14. protective	
Type	Collocation Errors and Corrections	Number (%)
	reserves (protected area); 15. deep accent (heavy accent); 16. weighty accent (heavy accent); 17. authentic accomplishment (real accomplishment); 18. invented account(fictitious account); 19. plugged drain(blocked drain); 20. whole accuracy (total accuracy); 21. occasional friendship (casual friendship); 22. a light acquaintance (a slight acquaintance); 23. supplementary activities (extracurricular activities); 24. intelligent acuity (mental acuity); 25. enduring address (permanent address); 26.thick admiration(deep admiration); 27. eyeless admiration(blind admiration); 28. choosy admission; 29. a fearless adventure(a bold adventure); 30. a dare adventurer (a bold adventurer); 31. a big adversary (a worthy adversary); 32. a foxy adversary (a subtle adversary); 33. wrong advertisement (false advertisement); 34. frank suggestion (frank advice); 35. present affairs (current affairs); 36. impermanent agreement (tentative agreement); 37. timeless illness (chronic illness); 38. subtle air (rarefied air); 39. rational absence(excused absence); 40. poor apartment (shabby apartment); 41. convinced appeal (an eloquent appeal); 42. insatiable appetite (unsatisfied appetite); 43. keen approval(heartly approval); 44. normal approval (official approval); 45. a national banquet (a state banquet); 46. coal monarch (coal magnate); 47 easy gains (easy pickings)	
L4 Quantifier + noun	1.a line of hope (a glimmer of hope); 2.a bundle of flowers (a bunch of flowers); 3.a cloud of wind (a gust of wind); 4. a little of humor (a bit of humor);	4 (2.4%)
L5 Adverb +adjective	1. definitely value (definitely worth); 2. strongly polluted(seriously polluted); 3. purely beautiful (flawlessly beautiful); 4. totally acquainted (completely acquainted); 5. greatly amusing(highly amusing); 6. very enjoy(enjoy greatly); 7. clearly absorbed(deeply absorbed); 8. deeply accomplished(highly accomplished).	8 (4.8%)

Type	Collocation Errors and Corrections	Number (%)
L6 Verb+ adverb	1. manage reasonably (manage effectively); 2. arrange properly(arrange systematically); 3. admit happily (admit readily); 4. fit wholly(fit perfectly); 5. act simply (act rashly); 6. celebrate gladly(celebrate joyously); 7. accept wholly(accept fully); 8. accept joyfully(readily); 9. accommodate quickly(accommodate easily); 10. report vividly(report directly); 11. express accurately(express clearly); 12. cheer readily (cheer enthusiastically); 13. act unlawfully(act illegally); 14. adapted fastly(adapted quickly); 15. adhere persistently (adhere consistently); 16. admire really(admire greatly); 17. admitted bluntly(admit readily); 18.; drive tiredly(drive carefully) 19. adopt accordantly (adopt unanimously); 20. advocate firmly (advocate strongly); 21. agree readily (agree completely) 22. amuse completely (amuse greatly); 23. acquaint totally (acquaint thoroughly); 24. nod mannerly (nod politely).	24 (14.4%)
In Total		167 (100%)



APPENDIX F

Writing Task 1 for RQ1

凯里学院_计算机科学与技术_分院 姓名: Li Deng
邮箱: 635---368@qq.com QQ 号: 635---368

Reduce Waste on Campus

As is known to all, Waste on campus has become a more and more serious problem^{L2}, we can easily see many students (F-free combination) dump a lot of food (F) in the garbage and some students spend thousands of Yuan buying fashionable clothes^{L2}. We have already wasted a great deal of precious resources^{L2} such as water and electricity and so on.

Our waste adds a burden^{<L1>} (impose a burden) to our parents and the society. The negative effects^{L2} of waste can be shown in the following aspects. In the first place, it makes some students dependent on their parent, which is harmful to their development. If they don't learn how to support themselves, they will be "useless people (F)" when they graduate. In the second place, we spend so much time^{L1} in playing, eating, dressing so we delay our study. However some still puzzle on how to eliminate waste ^{L1} on campus. Here are some suggestions(F). To begin with, we should spend more time^{L1} on studying secondly we should turn off the light^{L1} when we leave dorms(F) or classrooms. Thirdly we should be economical in our daily life^{L2}. For instance, when go shopping(F) we should buy only what we really need. The last but not the least, we should stop wasting water(F) when we bath or clean faces (F), or wash clothes^{L1} and dishes that is because there is a great shortage of water in the world.

Actually the solutions vary according to different situations(F). Therefore, I believe that we can reduce the waste on campus with our insistent efforts^{<L2>} (ceaseless efforts).

Writing Task 2 for RQ1

How I Finance My College Education

Nowadays, tuition and fees for college are much higher than ever before. How to finance our college education has become a matter of concern for many students, especially those who from poor area (F).

Here are some ways to solve the problem L1. For the first place, we can only get money (F) from our parents but we can apply for a loan especially set up for college students. Our daily expense (F) can be added. We can also acquire the money <L1> (earn money) entirely<L6>(acquire gradually) by ourselves, keeping a part-time jobL1 in our spare time L2 and doing a full-time jobL2 in summer and winter holidays(F).

As far as I am concerned, I prefer the last way, I don't want to depend entirelyL6 on my parents, which will be a burden to them. I do not want to be so busy working a full-time jobL2 that I haven't enough time L2 to study. A part-time jobL2 will make our college life rich and colorful. We can learn some applied knowledge <L2>(practical knowledge) and get much experience(F). It will make important function (F) for our future job(F).

Certainly, it is necessary for us that we shouldn't waste our moneyL1. It is very important L5 to finance our college Education.

APPENDIX G

Pre-writing in the Second Study

凯里学院_旅游发展_分院

姓名: Fan Zhou

邮箱: 537---366@qq.com

QQ 号: 537---366

Should students be allowed to live off campus in KU?

Many students want to live off campus in KU. Personally, I am in favor of students should not be allowed to live off campus in KU.

Firstly, I think students rent a house (F) outside is not safe. You know our school is far away the center of the city, there is may be have some people (F) who are not good. And students' safe may be influenced.

Secondly, if we rent a house (F) outside it will cost more money L1. Our parents give us some money(F) every month(F). We shouldn't add a burden<L1> (impose a burden) to our parents. We should take care of our parents.

Finally, we can't control ourselves. We are wholly acquainted <L5> (thoroughly/closely/intimately acquainted) after living together for a month. And sometimes we may play many games L1, watch movies L1 so late. Then we spend little time L1 in reading books L1 or doing homework L1. It's bad for our study. If we live in school we could not play so late.

So I just think that students should not live off campus. It is not suit us. I think it not good way (F) to living. As for students, we should live in school. That's all my opinion.

Post-writing in the Second Study

凯里学院_旅游发展_分院

姓名: Fan Zhou

邮箱: 537---366@qq.com

QQ 号: 537---366

Should students be allowed to live off campus in KU?

Some time ago, Teaching Affairs Office in Kaili University banned students from renting private accommodation (F/F). As for the question about students should be allowed to live off campus, different people (F) have different ideas (F). In my opinion, students should not be allowed to live off campus.

First, I think it is not safe that students live off campus. We know most of schools are located at developing area (F). And there are a few people so the students are dangerous.

Second, if we live off campus, we will take more time (F) to go home (F). So our study time will relatively lack <L5> (relatively short). For example, from our dormitory to our class, we can take 20 minutes. But we will take 30 minutes to the class when we live off campus.

Finally, I think the students can cost many money L1. When we live outside L6, we must pay other expenses (F/F). And it is not good for our life.

So, I think the students are not allowed to live off campus in KU.