國立政治大學英國語文學系碩士班碩士論文

指導教授: 張郇慧 博士

Advisor: Hsun-huei Chang, Ph.D

以語料庫為本之近似詞教學成效之研究:以台灣大學生為例

The Effect of Teaching Near-synonyms to Taiwan EFL University Students: A Corpus-based Approach

研究生: 陳聖其 撰

Name : Sheng-chi Chen

中華民國 103 年 1 月

January, 2014

The Effect of Teaching Near-synonyms to Taiwan EFL University Students: A Corpus-Based Approach

A Master Thesis

Presented to

Department of English,

National Chengchi University

In Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

by

Sheng-chi Chen

January, 2014

The members of the Committee approve the thesis of Sheng-Chi Chen defended on January 13th, 2014

Hsun-huei Chang

Professor Directing Thesis

One-Soon Her

Committee Member

Shu-hui Chen

Committee Member

Approved:

Tsui-fen Jiang, Chair, Department of English

ACKNOWLEDGEMENTS

I am very grateful to many people that have contributed in indirect or direct ways to this thesis. First, I would like to express my gratitude to my dear advisor, Dr. Hsun-huei Chang for her patience, encouragement, and guidance throughout the research process. During the thesis writing, she provided me with clear directions and revised my thesis. I really appreciate what she has done for me as an inspiring teacher.

Gratitude is also extended to my committee members, Dr. One-Soon Her and Dr. Shu-hui Chen, for providing me with valuable suggestions to help me improve the quality of the thesis. In addition, my special thanks go to my senior Claire who helps find the participants in my research. This study would not be completed without their help.

Fourth, I would like to express my indebtedness to Professor Siaw-fong Chung. I learned a lot of corpus knowledge and techniques from her courses. Her work in corpus linguistics and second vocabulary language learning inspired me for this thesis. I owe most of my skills as a researcher to her.

Next, I would like to thank my classmates in NTUE and NCCU. With their generous assistance and encouragement, my graduate school life was enriched. Moreover, I also want to express my gratitude to my friends, Winona and Zack, who help me check my thesis grammars and sentences.

Finally, my deepest gratitude goes to those whose influence may not be so evident but who nevertheless provided the educational and emotional foundation without which this thesis would have never been possible: my parents. Their love and faith gave me power and strength to go through every difficult time. I love you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

國立政治大學英國語文學系碩士班

碩士論文提要

論文名稱：以語料庫為本之近似詞教學成效之研究: 以台灣大學生為例

指導教授：張郇慧博士

研究生：陳聖其

論文提要內容：

台灣英語教育多以考試取向為主，許多教師進行英語字彙指導時採用填鴨式教學，致使學生無法於新的情境靈活使用字彙。

本研究旨在於探究以語料庫為本之教學對於台灣大學生在英語近似詞學習成效的影響，以台北市某一所大學 86 位英語學習背景及能力相似之大一生為研究對象。研究人數均分成兩班進行教學實驗，一班為實驗組，以資料觀察法進行教學，另一班為對照組，以傳統形式教學為主，每週一次五十分鐘，共進行十週。資料蒐集包含近似詞學習成就測驗前、後測，並且依據研究對象於實驗教學結束後接受語料觀察教學法回饋問卷，蒐集研究對象對於語料觀察法之反應與感知，進行量化分析。最後，透過訪談高分組和低分組學生，蒐集其質性資料進行研究探討哪些因素會影響不同英語能力學生對於資料觀察法的意願與需求。本研究發現如下:

一、近似詞教學有助於提升台灣大學生的英語字彙能力。兩組教學均在後測有

進步。但就後測成績來說，實驗組顯著優於控制組。資料觀察法之近似詞教學均較傳統教學法更能有效提升學生的英語字彙能力。

二、在不同程度的學生學習成效上，高、低分組學生均在後測成績有進步。對於高分組而言，實驗組後測成績顯著優於控制組後測。但對於控制組而言，實驗組的與控制組的後測成績未呈顯著差異。

三、大部分的學生對於運用資料觀察法學習單字均給予正面回饋，也肯定資料觀察學習法活動的效益。另外，根據高、低分組學生訪談結果發現，英語程度的高低的確會影響學生對於資料觀察法的喜愛和需求。高分組的學生希望先以資料觀察學習法為開端，再以傳統講解式方式做總結。但對低分組的學生而言，喜歡參與小組討論。由於單字量的不足，低分組學生希望在語料庫為主的教材旁能附上中文解釋，降低學習焦慮。

根據上述研究結果，本研究建議大學英語教師在教學現場能夠融入語料觀察學習法並依照不同程度的學生進行教材設計，以助提升學生學習英語單字。

關鍵字:資料觀察學習法、近似詞、語料庫為本

ABSTRACT

Corpus Linguistics has progressively become the center in different domains of language research. With such development of large corpora, the potential applications and possibilities of corpora in second language teaching and learning are extended. A discovery-based authentic learning environment is provided as well as created by such corpus-based language learning. Synonym or near-synonym learning is a difficult aspect of vocabulary learning, but a linguistic phenomenon with ubiquity. Hence, this research aims to investigate the effectiveness of the application of data-driven learning (DDL) approach in near-synonyms instruction and compare the teaching effect on the high and low achievers through the near-synonyms instruction.

Participants of this study were given instruction throughout the eight-week corpus-based teaching with materials compiled by the teacher. This is a quasi-experimental study consisting of comparison between two experimental conditions, with a pre-post test and control-experimental group design, followed by qualitative method of semi-structure interviews and questionnaire provided to the experimental group of EFL university students in Taiwan. Two intact classes of 86 college students participated in this study. The quantitative analysis of the pre- and

posttest scores and questionnaire were conducted through descriptive statistics and frequency analysis in order to explain the learning effects and learners' perceptions.

The results of the study revealed that: (1) participants in the experimental group made significant improvement in the posttest; (2) EFL high proficiency learners with DDL approach performed better than high achievers who were taught by the traditional method. However, low achievers may not be able to benefit from DDL approach in the form of concordance teaching materials; (3) the majority of the participants had positive feedback on DDL activities. Also, types of preferred DDL activities were strongly influenced by students' proficiency level. Low achievers preferred activities that should involve Chinese translation as the supplementary note while as for the high achievers, they were looking for the teacher's explanation of words' usages and functions in the end.

This study demonstrates the importance in illuminating the dynamic relationship between DDL approach and second language near-synonyms learning, as well as provides English EFL teachers with a better concept to incorporate corpus or concordance lines into vocabulary instruction.

Key words: data-driven Learning, near-synonym, corpus-based approach

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Taiwan is a test-oriented society, where many EFL learners are taught to learn English vocabulary through drilling and memorization because this is thought to be the best way to improve learners' vocabulary. Although this method seems to be an effective way to cope with tests in school, students who undergo the school entrance exam system in Taiwan might become confused and have problems in using precise vocabulary when dealing with new contexts.

Nevertheless, when students encounter a new vocabulary, teachers, before they think of the best way to explain or explore vocabulary, tend to use the traditional 'drills and repetition', methods to make students memorize vocabulary. Due to the fact that the majority of junior or senior high school teachers are under the pressure of limited time and needing to maintain students' academic performance, few teachers try to find a new way in helping students understand the usage. This is the situation in Taiwan, a test-oriented, EFL context, in which rote-learning is the most common method of learning vocabulary.

When the students enter colleges, the teaching context will be changed with a more flexible surrounding. Students have no test pressure as previous learning stages. The teacher in college might have flexibility to arrange his or her own curriculum and to teach students in different ways. This study focuses on college students because they have no test pressure and more time to explore the usage of each word. For college teachers,

although the class still has limitation to some extent, teachers can still manage to do a lot of creative activities in class.

In addition to the need of vocabulary development, the strategy of how to deal with bimodal distribution phenomenon of students' English proficiency level in Taiwan EFL context has become an important issue for teachers.The performance gap in English proficiency has troubled a lot of English teachers to meet the needs of all students. The effectiveness of vocabulary acquisition lies much on learners' ability to produce or recognize a word in different contexts, which requires the retrieval of the vocabulary knowledge in long-term memory (Craik & Lockart, 1972). Hence, when teachers want to design teaching materials, they need to first understand the learning distinctions between high and low achievers. Teachers will be able to respond to both groups' needs towards learning vocabulary.

With the rapid development of large corpora and the availability of powerful personal computers, many researchers and teachers in ELT have begun to explore the potentials of corpus linguistics in language teaching and language learning. One of the ways to learn vocabulary is using a concordancer to learn vocabulary. Concordancer is not only a sophisticated computer retrieval program with a large amount of information in the form of computer language corpora facilitating data-based inductive learning (Chan & Liou, 2005), but also a bundle of examples of a specific word or phrase that can be displayed efficiently. Thus, the amalgamation of corpora and concordances provide a new framework in the language learning and teaching by allowing learners to discover patterns through extensive naturally occurring examples in real texts (Hill, 2000). Several scholars have conducted the use of concordancing for second or foreign language acquisition, such as collocation learning (Lin, 2002; Tseng, 2002; Chan & Liou, 2005), lexical acquisition (Cobb, 1999; Sun & Wang,

2003; Supatronant, 2005; Yeh, Liou & Li, 2007; Boulton, 2009), and grammar acquisition (Hsieh, 2008; Ewa, 2011; Wang, 2012).

Data-driven learning (DDL) exploits the computer corpus for teaching and is developed by Tim Johns (1991), a pioneering teacher who combined corpus with language teaching (Leech, 1994; Hunston, 2002). The essence of DDL, a *research-then-theory* method, is that students acquire grammar rules and regular patterns inductively through exploring corpus materials. In the DDL approach, based on the theory that students can act as 'language detectives (Johns, 1997:101),' learners are not regarded simply as recipients of knowledge, but as researchers studying the regularity of the language. Teachers present learners with authentic examples and ask them to interpret clues and draw conclusions from context and this encourages learners to 'search for information without knowing in advance what patterns they will discover (Huston, 2002:184)'.

With the large interest in data-driven learning (DDL), a growing number of studies have suggested that data-driven learning may indeed facilitate processes beneficially to second language learning, especially to develop learners' writing and vocabulary ability (Gaskell& Cobb, 2004; Scott& Tribble, 2006; Boulton, 2009) to name but a few. More Specifically, both low and high proficiency level students increased their writing ability (Yoon& Hirvela, 2004; Tian, 2005; Boulton, 2009) and correct usage of grammar increased (Ciezielska-Ciupek, 2001; Tian, 2005). Students also view data-driven learning as a conducive method, which can result in their progressively accurate usage of vocabulary and grammar (Nesselhauf, 2003; Chen, 2004; Koosha & Jafarpour, 2006). Wu (2010) has discussed the effect of integrating corpus data into grammar instruction by using three near-synonyms 'advise', 'recommend', and 'suggest' as examples. Near-synonyms, which can be defined as

lexical pairs that have very similar cognitive and denotational meanings, but which may differ in collocational or prosodic behavior, such as ‘strong’ and ‘power ’(Halliday, 1976). However, the issue of teaching near-synonyms has been totally ignored in previous literature. In order to help fill this gap, this study establishes a niche to investigate the possible existence of using DDL approach to teach learners’ near-synonym and also investigate their collocational competence. The goal of the study is to find out whether the learner knows the correct usage of near-synonyms and their phrases expressions through the DDL approach.

1.2 Significance of the Study

Data-driven learning (DDL) is really a “brand-new” teaching methodology in Taiwan with regard to the EFL teaching context. In Taiwan, the majority of data-driven related studies were conducted in four domains: vocabulary learning and teaching, collocational acquisition, second language grammar learning, and second language writing. For EFL students, vocabulary building is really important for their academic success. However, most literature and studies related to the topic of vocabulary have emphasized the effects of learners’ collocational competence on vocabulary learning regardless of students’ proficiency level, and little research has been explored on the vocabulary issues of ‘near-synonyms’. The current study aims to explore the effectiveness of the application of data-driven approach in near-synonyms instruction. A secondary purpose of this study is to compare the teaching effect on the high and low achievers through the teaching of near-synonyms.

The expected benefits of this research would show that data-driven learning of this research could be helpful not only for university students in a mix-competency

classes but also for college teachers on the front line who are considering incorporating corpus or concordance lines into their vocabulary instruction.

國立政治大學
National Chengchi University

CHAPTER 2

LITERATURE REVIEW

This chapter is divided into three sections. The first section elaborates how data-driven learning (DDL) is applied in classrooms, using consciousness-raising activities and concordance lines as the teacher-made teaching materials. Secondly, some studies related to the impact of data-driven learning and teaching also reviewed in terms of three teaching perspectives: vocabulary, collocation, and grammar learning. In the third section, some local and international studies regarding synonym and near-synonym applied in English learning and teaching are explored.

2.1 Data-driven learning (DDL) in a classroom

Applied Linguists working on technology-related issues have for some time noted the relevance of technological changes in the digital global economy for TESOL (e.g., Cummins, 2000; Warschaucer, 2000; Chapelle 2001). In the area of language teaching and learning, the potential applications of electronic corpora, the electronic authentic language databases available on the Internet or stored in personal computers (Serkan, 2011), in language teaching and learning have receiveconsiderable attention in recent years. It is also possible for teachers and learners to access corpora by themselves (Gaskell& Cobb, 2004; Scott& Tribble, 2006; Boulton, 2010) to name but a few.

The English language teaching (ELT) profession is currently undergoing what has been labeled by some as a major paradigm shift (Woodward, 1996). Starting from the mid 1980's (Swain, 1985), some began to question many aspects of Communicative Language

Teaching (CLT). Since then, many more have gotten together to express the idea concerned with the direction that the position of CLT will be taken. Later on, the idea of DDL was proposed by Johns (1991) as an innovative approach to the implementation of concordancing materials, which lists all the occurrences of a word in a text with enough co-text in each line in the field of second language acquisition (SLA). The readers can tell how the word is used and where it was used. According to Johns (1994), the language learner using this approach is essentially a research worker whose learning is driven by access to authentic linguistic data.

Batstone (1995) claims that DDL is a pedagogic continuum (see figure 2.1) from product to process. It has the advantage of product approach since the specific aspects of language are presented to the learners by multiple exposures within contexts. Simultaneously, it offers a completed process approach because DDL promotes creativity and self-discovery learning among learners. While DDL is a very new methodology, DDL appears to utilize the strengths of both product and process approaches to teaching grammar and collocation successfully (Hadley, 2004).

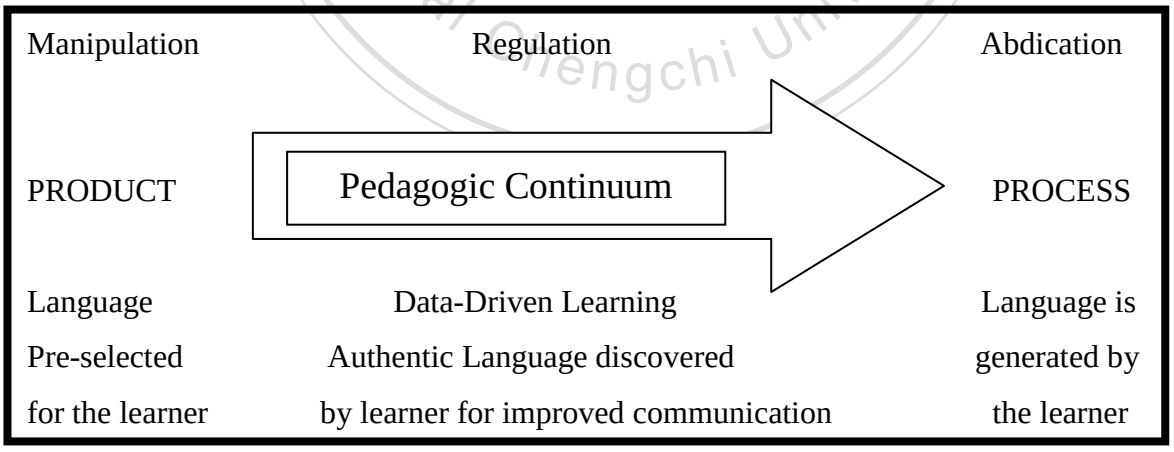

Figure 2.1 A pedagogic continuum from product to process grammar learning through DDL. Adapted from Batstone (1995).

DDL approach suggests that grammar and collocation learning should consist

largely of consciousness-raising activities rather than teaching rules (Willis & Willis, 1996). Consciousness-raising (CR) is defined by Rutherford and Smith (1988:108) as "the deliberate attempt to draw the learner's attention specifically to the formal properties of the target language." Kennedy (2003) also holds the similar view, and explains that the second vocabulary acquisition process as depending on the consciousness raising and learners should be exposed to authentic language materials as much as possible. Moreover, Schmidt (1990) argues that it is necessary to 'notice' features of the target language at the beginning in learning process. In Ellis' (2003:163) study, he also claims that 'consciousness-raising constitutes an approach to grammar teaching which is compatible with current thinking about how learners acquire L2 grammar.' If the teacher can place students in the position of researchers, this will increase learners' self-esteem and confidence (Johns, 1994; Willis & Willis, 1996; Johns, 2002).

Data-driven learning, a "research-then-theory" method which can be used to study grammar and collocation, is just like the consciousness-raising activity in that language learners start with a question, and then come to their conclusions after analyzing the corpora with a concordance program. Hence, the teaching material has become an important element when the teacher wants to cultivate learners' linguistic awareness. Recently, scholars who support DDL have written an increasing amount of material to show how data from corpora and concordances can be used in the classroom. The results all indicate that learners have a big improvement after corpus-based teaching instructions (Tribble & Jones, 1990; John, 1994; Kettemann, 1995; Tribble, 1996). In these teaching materials, concordance plays an important role in spite of the different teaching. Teaching and learning purpose is to learn grammar patterns and collocations through the concordance line. More research will be necessary to determine if learning actually takes place from a DDL approach. If it could be shown the DDL approach is not only engaging

but also effective in language learning, it would be a major step toward establishing it in Taiwan EFL classroom. This study uses data-driven learning to investigate the learning effectiveness of learners.

There is a controversial issue in the relationship between data-driven learning and learners' language proficiency. Some scholars think DDL is helpful only for high achievers in learning English vocabulary, while some claim DDL helps low achievers too. For high achievers, the proponents of DDL take it as a task-based learning which requires good language awareness to find patterns and rules of teaching materials presented to them (Touraj & Zahra, 2012). Thus, it is claimed that DDL is only useful for the learners at advanced levels and a lot of training is essential for its application. Johns (1986) pointed out that DDL is appropriate for the learners who are adults and have high motivation and intelligence. Similarly, in 2008, Boulton surveyed fifty empirical DDL studies and noted that only four were conducted with beginning and low level studies. A lot of research on DDL reveals that this type of learning has been used for the learners at advanced level. Boulton (2008:39) also noted that, "current research encourages the belief that DDL is only useful for advanced learners in a computer laboratory, and with the researchers devoting considerable time to developing corpora and training learners in small groups". The reason why most previous DDL studies focus on high proficiency learners might be that the participants were chosen according to the researchers' preference.

Although there is a tendency towards thinking that DDL is only useful for the learners at advanced proficiency level, some investigations present evidence that DDL is helpful for the low proficiency level students as well. Sealy and Thompson (2007) proved that even primary school students can take advantage of corpora in their native language without a higher order thinking skills. However, Braun (2007) pointed out that when

corpus is used for learning in ESL or EFL context, the situation is totally different from using DDL to learn first language. It might become less efficient. This is the reason why in most articles about DDL, the participants are chosen from high proficiency level. Since very little DDL research ventures on low achievers in EFL learning context, this present study tries to investigate the effect of using corpus-based DDL on low achievers' synonymous vocabulary learning, compared to traditional teaching.

2.2 Concordance lines and Data-driven learning

Concordance lines have been used as the teaching material in the classroom since the 1980s as part of the "data-driven learning" (DDL) proposed by Tim Johns (Sinclair, 1991). It is widely recognized as providing a fast and powerful means for learners to become more aware of the language and the patterns of its various grammatical features, words, and lexical patterns (Tribble & Jones, 1990; Sinclair, 1991; Mindt, 1997; Fox, 1998; J. Willis, 1998; Honston, 2002; O'Keefe et al., 2007). An important effect of concordance lines is that they enable learners to take a more objective look at the language. They are like 'tiny snapshots' (Willis, 1998) of a linguistic landscape. Hunston (2002: 170) also noted that, "DDL involves setting up situations in which students can answer questions about language by studying corpus data in the form of concordance lines or sentences".

Johns (1997) points out that a concordance can bring to the class abundant authentic language examples that can be studied and exploited in many ways by learners, so that the learners can see how the target word collocates with other words, which grammar patterns will be followed, which preposition the target word goes with, and so on (Willis, 1990).

Furthermore, using concordance line as the teaching material may be a more

learner-centered methodology. Johns (1994:101) refers to learners who are studying by concordance line as so-called “ language detectives ” whose task is to discover the rules or patterns of the language they are focusing on by finding, identifying and inferring these linguistic implication from the context (Serkan, 2011). Learners can be motivated to discover new meanings and to examine collocations because language learning is more likely to occur when adequate examples are noticed and processed by learners (Cobb, 2003).

As for learner motivation, Tribble and Jones (1990) claim that concordance line can activate learner motivation. They believe that “ the fact that the source material for exercises is drawn from real life rather than concocted by teachers to increase motivation, as it gives learners immediate contact with the target language in use” (p. 38). However, Tomlinson (1990) warns instructors against assuming that all of their students are “ready and willing” to learn any particular teaching points chosen for presentation. Though it proves to be an effective teaching methodology, many people do not believe concordance lines are suitable for lower or beginning level students, especially for weak and remedial beginning students (Willis, 1998). The teacher has to compile and tailor the concordance lines before actual teaching in the classroom. Manual concordance lines have become an important teaching material when the teacher wants to teach through corpus-based approach. Figure 2.2 shows an example of the manual concordance lines of the target word ‘people’. Learners can read the context with each target word in each concordance line and then generate the usage and collocation by themselves.

1	a three-month contract with ACET, is looking forward to caring AIDS. Letters to the Editor Letters	**people** with for HIV/
2	To reduce the number of new HIV infections by giving young AIDS. AIDS Care Education	**people** the facts about
3	, under the auspices of the United Nations. At least nine between November 1990 and January 1991 by death squads linked	**people** were killed
4	may be POCs. Among the possible prisoners of conscience are 37 up to 20 years' imprisonment for their involvement in	**people** sentenced to
5	justice. PRISONER LETTER WRITING CAMPAIGN. Each of the follows is a prisoner of conscience. Each has been	**people** whose story

Figure 2.2 *Example 'people' for manual concordance lines (extracted from BNC web Corpus)*

With manual concordances, learners will have the chance to not only 'notice' the language in each text they encounter, but also 'recognize' the word they have seen it before (Allen, 2012). The rationale of manual concordance lines is like the 'spiral structure' in the textbook development. The intertexual nature of course books will also allow learners who choose to develop it to break through a new language awareness and increase the sharpness of their learning curve. Since, a good general English course book will recycle important language from one lesson to another; manual concordance lines can serve as the same purpose. As Willis (1998:63) observes that manual concordance lines, as a learner-centered methodology, is ideal for mixed level class, since it allows students to work at their own level, in their own time and in their own way. Manual concordance will not only help learners process new language more effectively, but also produce the language repertoire more easily.

2.3 Related studies on data-driven learning

With the emergence of data-driven learning, the central task of the language learner changes. Rather than dull memorizing grammatical rules and word lists which lack context, learners are more focusing on learning the lexical chunks (e.g. collocation, discourse) and pragmatic convention that shapes the use of the lexis (Tribble & Jones, 1990; Lewis, 1993; Fox, 1998; Willis, 1998; O'Keefe et al., 2007). With rich 'contexts', generally speaking, data-driven learning can be conducted in three realms: vocabulary, collocation and grammar learning and teaching.

2.3.1 Vocabulary

Cobb's (1999) study found that the experiment group who learned language through concordance has a higher proportion in their retention than the control group, a dictionary-based approach. Boulton also (2009) concluded that the experiment group who learned adverbs from corpus has better retention than the control group who was taught by traditional teaching methodology. Yeh, Liou and Li (2007) conducted a synonym research by means of using data-driven approach in Taiwan colleges. They tried to find whether the learner can use synonyms effectively in their writing. For example, the students may substitute *great*, *large* or *huge* for *big*. The result showed that the scores of post-test are higher than pre-tests, but, the delayed post-tests score is lower slightly than post-tests. All in all, the students writing ability improved significantly after conducting the teaching experiment. Similar results were reported in Supatronant (2005) study. The study elaborated on the purpose of how concordance output software is in direct relation to definitive and productive knowledge of learners at the engineering department. The findings reported that students achieved greater progress in operating concordance output software.

2.3.2 Collocation

The importance of collocations has long been stressed by scholars involved in teaching second or foreign languages (Xiao & McEnery, 2006). In second language learning, the knowledge of collocation effectively helps learners to generate sentences, to attain proficiency in writing, and to become fluent, native-like L2 speakers (Zang, 1993). On Taiwanese learners, Lin (2002) examined high-achievers' and low-achievers' performance in receptive and productive verb-noun collocation tasks after six weeks collocation instruction. Thirty-two students were divided into two groups. One class was conducted with the deductive teaching approach and the other with inductive approach. The result pointed out all students made more progress in receptive tests than productive ones, and high-achievers generally performed better than low-achievers. The result suggested that learner's competence might have impact on the effect of data-driven learning. Chen and Liou (2005) investigated the influence of using data-driven learning with a bilingual concordancer on English verb-noun collocation learning. The study revealed that students had more progress than their pre-test scores. In that, it is beneficial to use data-driven learning approach to teach collocation (Wu, 1996; Liou, 1999 a, 1999 b; Chen, 2002; Lin, 2002; Liu, 2002; Chen & Liou, 2005; Li, 2005; Shin & Wang, 2006).

2.3.3 Grammar

In EFL context, an important feature for developing students' grammatical competence is to provide them with a sufficient amount of language samples. Researchers such as Dekeyser (1995) and Doughty and Varela (1998) have all suggested that grammatical correctness in communication could be achieved by the analysis of forms and of functions with a bundle of examples. This has been one of the reasons that support the teaching rationale to use a corpus-based material as an example of the natural

resources.

In Wang's (2012) study, she investigated whether data-driven learning, as compared to focus on forms of grammar rules, is helpful for 5th grade English language learners in Taiwan in understanding and acquiring the English grammar points. The result indicates that the data-driven learning group performed better than the focus on forms learning group on the 'third-person singular'. Ewa (2011) showed how competence in the understanding and creation of discourse can be achieved by the use of a corpus concordance and by the students performing various language-learning tasks. The concordance and the related task were assumed to be useful tools in developing students' competence in formulating appropriateness in discourse. The findings showed that a corpus-based approach not only raised students' grammar awareness, but also guided the students' attempts at 'rule discovery'. Hsieh (2008) provided a syllabus which was based on corpus-based grammar syllabus design with three near-synonyms 'see, watch, look at'. She suggested that grammar should be learnt with the context and a small number of concordance line is not sufficient.

2.4 Synonym and near-synonym in language learning and teaching

Vocabulary is the most common subject taught directly by teacher through corpora. Among vocabulary, synonym or near-synonyms are the most difficult vocabulary to learn (Higa, 1963; Lanfer, 1990; Tsui, 2005; Liu, 2010). Tsui's (2005) research points out that for ESL teachers, one challenge in vocabulary instruction concerns the synonyms and near-synonyms. Synonym or near-synonym is a common yet complex linguistic phenomenon with its ubiquity in language. Liu (2010) assumed that while synonyms express basically the same concept, they often do so in different fashions, for different contexts, and/or from different perspectives. In other words, synonyms are often not identical in meaning and hence not completely interchangeable. Therefore, synonyms are

a challenging and simultaneously important lexical category because they are essential for expressing shades of meaning to help us convey our ideas and feelings accurately for effective communication (Edmonds & Hirst, 2002; Taylor, 2003; Divjak, 2006; Divjak & Gries, 2006). In linguistic term, they "are neither in free variation, nor in complementary distribution" (Divjak, 2006). Cruse (1986) remarked that the synonym relationship seems to be unstable, in that one of the synonyms, being dysfunctional, will be liable to fall into disuse; alternatively some semantic or stylistic nuance will in time come to differentiate words. Bolinger asserted (1977) that "if two ways of saying something differ in their words or their arrangement they will also differ in meaning. Similarly, as Room (1981) mentioned in his paper, the distinction between *forest* and *woods* is a complex combination of size, proximity to civilization, and wildness as determined by the type of animals and plants therein. Research near-synonyms includes *'little/small, high/tall, start/begin, stop/finish, motor/engine'* (Taylor, 2003), *cut/break* (Taylor, 2007), *fangbian 'convenient'* and *bianli 'convenient'* (Chief, Huang, Chen, Tsai, & Chang, 2000), *gan3* 趕 *'rush'* and *qiang3* 搶 *'rush'* (Chung, 2008) and many more.

However, only limited research has investigated the effects of synonymy on vocabulary learning, especially by means of data-driven learning, a corpus-based approach. Higa (1963) found that learning two synonyms at one time is more difficult than learning two unrelated words. Laufer (1990) also found that synonymy is one of seven interlexical factors that can reduce the chances of vocabulary acquisition. She gave two reasons that synonyms may be more difficult to learn than other words. First, learners often make mistakes using synonyms because some of them may be substituted effectively in some contexts but not in others. For example, *strong* and *powerful* have similar meanings, but usually tea is only *strong*, and engines may be *powerful* but are rarely *strong*. This is an important point and demonstrates that some synonyms might be

more difficult to learn than others because synonyms with similar meanings do not always have the same collocates. Certainly, some words such as *good* and *nice* are synonymous in many contexts, while others such as *powerful* and *strong* are synonymous to a lesser degree. The second reason she gave is that less advanced learners are unlikely to try to learn words with similar meanings when they have a greater need to learn unknown L2 meanings. Since near-synonyms are hard to learn, teachers might put more emphasis on them when teaching vocabulary.

2.5 Chapter Summary

To summarize, vocabulary is important for any language learning. For most EFL learners, learning vocabulary requires time and efforts. When it comes to classroom instruction, language teachers should include activities that facilitate vocabulary learning process. In Taiwan EFL context, teachers often use the traditional deductive methodology, they might face the common problem of ‘how to give students correct and practical language usages (Conrad, 2000)’. This is still a controversial issue in second language learning and teaching.

Then the corpora teaching methodology ‘DDL’ proposed by Tim Johns, is the key methodology to teaching learners through corpora. Later, more and more researchers dedicated themselves to the research of corporal-based teaching method (Willis and Willis, 1996; Hunston, 1998; Cobb, 2003; Sun and Wang, 2003; Yeh, Liou & Li, 2007; Hsieh, 2008). They point out that through data-driven learning (DDL), students can not only learn the correct target language functions and usages from authentic materials but also acquire the vocabulary knowledge and grammar knowledge from concordance lines. Students might feel learning vocabulary is not a dull and abstract curriculum. Taiwanese EFL teachers still have doubts about the effectiveness of DDL. According to the literature,

DDL can be conducted in three realms: vocabulary, collocation and grammar learning and teaching.

Identifying and classifying near-synonyms are a big challenge for second language learners because teachers seldom emphasize this part. Hence, it is necessary to teach near-synonyms in class. There is very little research on the effects of synonymy on vocabulary learning, especially by means of data-driven learning. Hence, the purpose of the present study is to explore the effect of using corpus-based DDL on learners' synonymous vocabulary learning, compared to traditional teaching (with dictionaries). It also aims at focusing on the learners' reactions toward the DDL approach.

2.6 Research Purposes and Research Questions

The study adopted two groups, experimental group with data-driven learning approach and control group with traditional teaching method. Participants received 10 weeks of data-driven learning instruction and traditional teaching methodology separately and the study would use both qualitative and quantitative methodologies. The present study is designed to address the following research questions that have not been answered in previous studies:

1. Does the data-driven learning (DDL) approach with manual concordances improve better in students' near-synonyms knowledge and performance in comparison to control group taught in traditional method?
2. Are there any significant differences between high and low achievers in the experiment and the control group?
3. What are learners' reactions and preferences regarding using the corpus to learn vocabulary?

國立政治大學
National Chengchi University

CHAPTER 3

METHODOLOGY

The researcher had a pilot study to see whether the DDL method and questionnaire are flexible or not. The pilot study is described first and then the quasi-experiment is introduced. The pilot study aims to investigate the influence of data-driven learning on three near-synonyms ‘people, person and human’ through corpus-based activity. There were three research questions in the pilot study similar to the quasi-experiment: (1) Does the data-driven learning (DDL) approach with manual concordances improve students’ knowledge and performance of near-synonyms? If so, is it only effective after a certain period of time? (2) What and which aspect of the three near-synonyms do the students have to notice? (3) What are the learners’ reactions and preferences regarding using the corpus to learn vocabulary?

First, corpus technique was put to use in one first-year EFL reading and writing class I for 13 English major students at National Chengchi University in the spring semester of 2012 academic year. Second, pre-test, post-test and delayed-posttest were conducted in order to compare the effectiveness of learners’ vocabulary knowledge. The main purpose of implementing this pilot study was to know whether students’word knowledge improved or not on three near-synonyms idiomatic expressions and the word usages. All the students’ native language was Chinese. Most had been learning English for 12 years at school.

In the Activity, the researcher put more stress on learners’ vocabulary learning strategies which based on the Schmitt’s (2012:207-208) taxonomy. The lesson plan in

pilot study is displayed in Appendix A. The activity in quasi-experiment was followed by the same procedure in pilot study. For teaching materials, teacher gave every student a reading package. The reading package consisted of a lyric (*where is the love*: The Black Eyed Peas) and seven pages concordance handouts. The lyric was used as a warm-up activity to activate students' motivation. Concordance lines were extracted and selected with careful consideration from the BNC web corpus and then compiled into four page handouts. The four-page concordance handout involved seven target 'phrases' and one 'idiomatic expression' which were based on three near-synonym (see Appendix B). Seven phrases and one idiomatic expression are shown in Table 1.

Table 3.1 *Idioms and Phrases in Three Near-synonyms*

	Phrase	Idiom
people	*boat people*	
person	*lay person, people person, living person, in the person of, be on one's person*	*snake oil person*
human	*complete human, human person*	

The results implied that participants made great progress in vocabulary knowledge in immediate posttest after data-driven learning instruction. In addition, the participants' delayed posttest values are still higher than pretest. This suggests that DDL (data-driven learning) can improve students' near-synonymous vocabulary knowledge. In addition, the result also suggests that data-driven learning might not improve participants' word retention ability. Such a finding echoes the previous studies (Sun &Wang, 2003; Chan & Liou , 2005; Koosha & Jafarpour, 2006; Yeh, Liou & Li, 2007; Alex, 2009; Ewa, 2011; Serkan, 2011) and proves that data-driven learning enhances participants' second language vocabulary acquisition.

For the learners' perception, learners' perceptions towards computer learning were

very positive. Learners thought that doing vocabulary exercises through paper-printed concordance is more interesting than doing identical exercises of traditional teaching from textbooks. Every learner would like to do other similar corpus activities in later classes and was satisfied with the experience of learning English vocabulary in a corpus-based learning environment. More students had a sense of achievement in the corpus-based data-driven learning class and more were satisfied with this method.

Since the pilot study showed that the learners have significantly improved on near-synonyms, the main study will try to enlarge the population and increase teaching hours to investigate whether the learners improve their word knowledge or not.

The present study intends to investigate whether learners' near-synonymous achievement is enhanced through corpus-based data-driven learning (DDL) instruction. In this chapter, detailed information about methodology is presented in the following order. It will begin with the introduction of the research design, followed by the participants and setting selected in this research. Later, specific procedure and materials will be outlined. Finally, the instruments will be described, both quantitative and qualitative methods will be applied.

3.1 Research Design

This is a quasi-experimental study with a pre-, and post- test control group design, accompanied with qualitative method of semi-structure interviews and questionnaire with the experimental group of EFL university students in Taiwan. The independent variable is the different teaching delivery modes (corpus-based data-driven learning vs. dictionary-based traditional vocabulary instruction) and the dependent variable is learners' vocabulary knowledge achievement. Control variables include learner's age, instruction hours, designed activities and previous learners' English proficiency. There are two

groups in this study: one experimental group and one control group. The experimental group received corpus-based data-driven learning while the control group received dictionary-based instruction in a traditional classroom. Posttests of near-synonyms were conducted in the 10th week. The detailed of the research structure is shown in Figure 3.1.

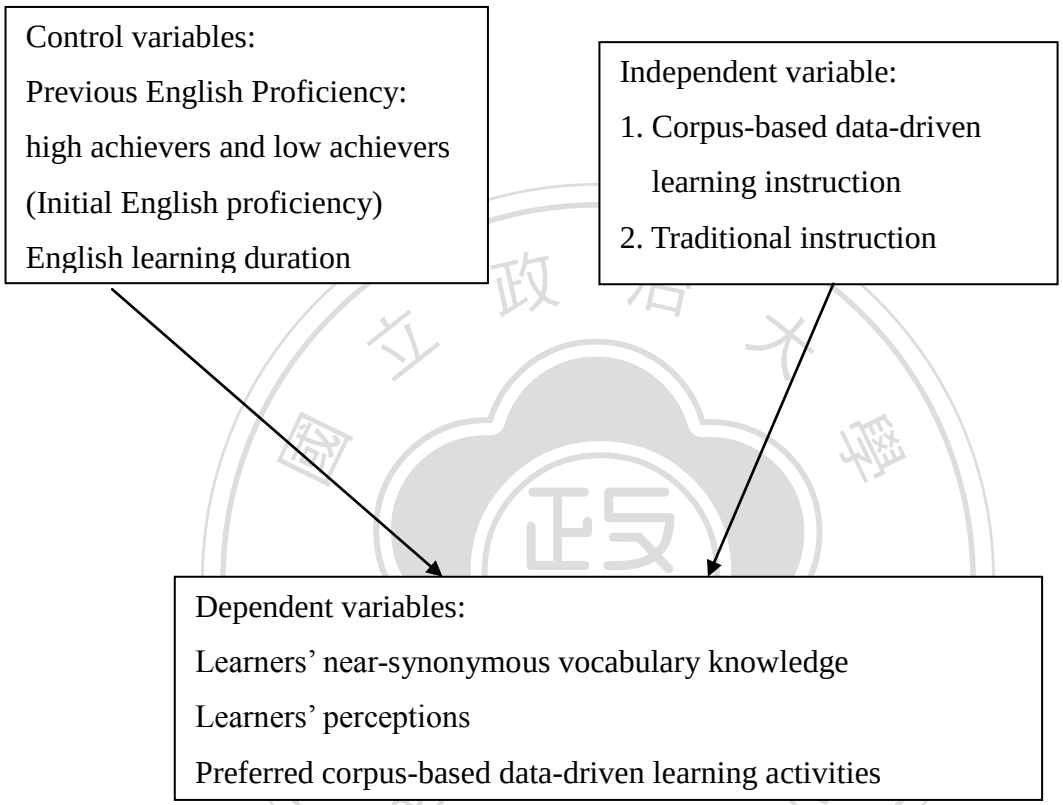

Figure 3.1 Research Structure

The variables involved in the study are explained as follows:

Control variables

In order to control the reliability of the study, several significant factors were strictly controlled in the experimental teaching procedures.

Since the classes were selected randomly by the researcher, students' English proficiency might be different from one another. The experimental group and control group had similar English proficiency test which was based on their General English

Proficiency Test (GEPT) intermediate level in the last semester of academic year 2013. The GEPT intermediate level test was used as a standard to group the participants in this study. Later, the researcher chose the highest and the lowest thirty percent of learners separately as the high and low achievers in the experimental and control group.

English learning duration

Duration means the time span the instructor gives lecture to learners. The instruction duration in both the experimental and the control group were the same. The instruction was given in each class for 50 minutes once a week, and it was given on 10 consecutive weeks. And the total instructional time is five hundred minutes.

Instructor

The instructor, who taught according to school schedule on Monday afternoon, was the same as researcher in this study. He remained the same attitude in order to avoid the so-called 'teacher's factor' interference, having no bias toward any particular students in the ongoing research. In the study, the researcher plays the role as a curriculum designer, activity facilitator, as well as an English instructor of near-synonymous vocabulary teaching.

Teaching materials

All students were taught near-synonyms in each class. The students in the experimental group received data-driven learning instruction based on the teacher-made concordance lines selected from the BNC *web* corpus. There were forty pairs of near-synonyms in the teaching materials. The control group was taught with the dictionary-based words' definitions which were extracted from the five on-line dictionaries (*Merriam Webster, Longman dictionary of contemporary English, Oxford*

dictionary, Cambridge dictionary, Collins English dictionary). Since both of on-line dictionaries were based on the word frequency, the teacher used these two dictionaries as the material to teach the control group.

Independent variable

This study adapted only one independent variable which was the teaching methodology related to data-driven learning approach. Students followed the same activity procedures in ten weeks.

Dependent variables

The dependent variables were referred to the students' performances which were affected by the instruction and their reactions toward the data-driven learning approach.

3.2 Research Procedure and Duration

The procedure of the study is shown in Figure 3.2 which was conducted through three phases, the preparation phase, the experimental instruction phase, and complete phase. In the preparation phase, data-driven learning (DDL) was determined to be implemented in the college freshmen English class to facilitate near-synonymous vocabulary learning. Simultaneously, the research questions were generated to help explore the effectiveness of DDL approach. Later, the researcher started to design activity and questionnaires as well as the format of pretest and posttest. After designing the whole research tool, a pilot study was conducted in National Chengchi University English major freshmen (*Reading and Writing Class I*) to determine if the research design was viable or not. After gathering the data from the pilot study, the research analyzed the results and revised the research tools to make sure the quasi-experiment can be conducted in a more

accurate way. Since participants in the two-year junior college had no displacement test, not like the majority of universities that divided students into different English proficiency level, it was necessary to identify the students level in two classes and then selected the high and low achievers. Hence, the researcher separated the students into two groups based on their GEPT test scores (intermediate level) before conducting the qusai-experiment. The students took the GEPT test in the last semester of 2012 academic year. The GEPT score was the reference for their initial English proficiency. The GEPT test includes forty-four multiple choice questions which consist of four parts—vocabulary and structure, fill-in-the blank and reading comprehension. The highest thirty percent of the participants was defined as high achievers, and the lowest thirty percent of the participants belonged to low achievers. Then the researcher divided the two classes into experimental and control group randomly.

During the experimental instruction phase, at first, a background questionnaire was delivered to survey students' learning experience and background information. In the first week, an introduction and an orientation were given in the first class. From the second class, data-driven learning approach was implemented in the experimental group while the traditional dictionary-based approach was the main instructing method in the control group in the vocabulary session in each English class. The experimental instruction continued for ten weeks, fifty minutes each time and once a week. After ten weeks instruction, students had a posttest to make sure if their near-synonymous vocabulary knowledge was improved. Later, a learner's perception questionnaire was given only in the experimental group to measure learners' affective skills toward data-driven approach. After filling the questionnaire, oral interviews were conducted to find out what effect data-driven learning instruction had on participants. Four of experimental groups, made up of high and low proficiency students, were served for the interview. The participants

were interviewed one by one. Finally, it was the data collection and analysis phase. After ten weeks, the data was analyzed through SPSS (17.0) independent *t-test* and pair-sample *t-test*. Later, according to the data results, the researcher wrote the thesis. The whole time table is presented in Table 3.1.

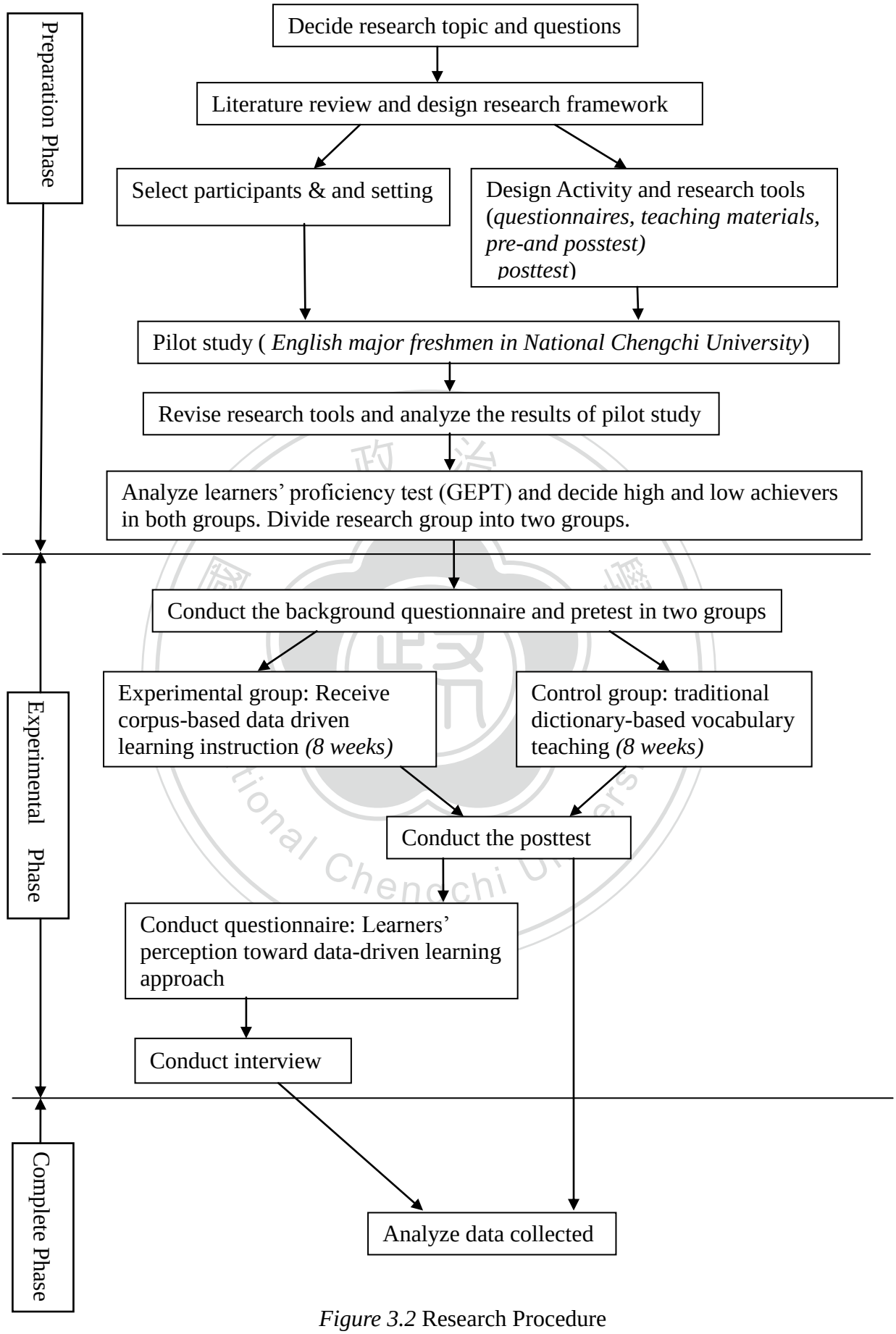

Figure 3.2 Research Procedure

Table 3.2 *Timetable of the study*

	Year	2012			2013								
	Month	10	11	12	1	2	3	4	5	6	7	8	9	10
Pilot Study Phase	Decide the research and questions	■	■	■										
	Design research framework and design Activity I in pseudo-class		■	■										
	Revise Activity I and design Activity II in pseudo class		■	■										
	Conduct research tools		■	■	■									
	Implement the pilot experimental instruction (include pre-and posttest)				■									
	Conduct delayed posttest				■									
	Analyzed collected data				■	■								
Quasi-experiment phase	Measure learners' initial language proficiency and group participants (high and low achievers)						■							
	Deliver pretests and background questionnaire						■							
	Conduct quasi-experiment instruction							■	■	■				
	Conduct posttest									■				
	Conduct perception questionnaire									■				
	Conduct interview									■	■			
	Data analysis and thesis writing										■	■	■	■
	Review Literature	■	■	■	■	■	■	■	■	■	■	■	■	■

3.3 Setting and Participants

Setting

This study was conducted in a College in Taipei City. This school is located in the center of a popular educational district, which is situated on the southernmost part of the Taipei City. There are five departments in this college. Students in this college have to study two years and then have a special test. There are totally 84 classes in the two-year college, 47 classes for the first-year, and 37 for the second-year. There are 44 students in one class. The enrollment is approximately two thousand students. Most students want to engage in government jobs in the future. Students can get familiar with English because they have a special examination in the second year of their college life. If they pass the examination, they will become a government official worker in Taiwan.

The classrooms in this college are equipped with a computer, a projector and a speaker. The teacher is able to access to the internet via overhead projector to display information from the websites or PowerPoint.

Participants

There are 47 classes in the first year in this school. Two English classes were selected randomly from all junior classes. Each class had 43 students (Female: 5, Male: 38). So, 86 students were participating in this experiment. Since the college did not have the English placement test, each English class was a mixed English proficiency class. The two classes were randomly assigned to the experimental group with concordance lines approach by using paper-based DDL material and the other was the control group with traditional grammar teaching respectively. According to previous studies mentioned in literature review, the results showed that DDL might have influenced on low and high achievers. Hence, in order to explore the influence of English proficiency on the

preference of data-driven learning, the 86 participants were further divided into high and low achievers according to their GEPT (intermediate level) scores in the previous semester. The researcher only discussed the highest and lowest thirty percent of learners in experiment and control group. Table 3.2 shows the distribution of participants in each subgroup.

Table 3.3 *Number of students in each subgroup*

Groups	Proficiency	Total
Experimental group	High	15
	Low	13
Control group	High	15
	Low	17

In each group, students were divided into eight subgroups, each group had five to six people. In each lesson, the teacher taught four pairs of near-synonyms, which meant the teacher taught at least eight words in one class. The words were selected from the 'vocabulary column' of learners' textbook and then the teacher added one or two words to form near-synonyms pairs. The format of data-driven learning activity is shown in Table 3.3.

Table 3.4 *Format of Data-driven learning activity*

	Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8
3/4	human person people	human person people	see watch look at	see watch look at	spend take cost	spend take cost	tour travel trip	tour travel trip
3/11	eventful Significant important	eventful Significant important	creative imaginative	creative imaginative	advantage benefit	advantage benefit	alter change	alter change

Two groups had the same pairs of near-synonyms and there were totally four pairs of near-synonyms in each lesson. In addition, every student was assigned a specific role, including group leader, timer or answerer, during the vocabulary exercise activity. Students were able to discuss the role distribution with their group members and decide each person's role. The students' language learning history was typical for Chinese. Most had been taught English for 12 years at school prior to entering university. The teacher used the BNC *web* corpus, a Web-based interface to the 100-million words British National Corpus (BNC), to teach experimental group. BNC*web* was developed for internal use at the University of Zurich (Lehmann et al. 2000) but was subsequently released to the general public (non-commercial use only) in the year 2002. None of the students had any prior experience of DDL.

3.4 Experimental instruction Design

3.4.1 Teaching Content

In this research, the researcher adopted participants' school textbook *Four Corners Level 4*, which was edited by Jack C. Richards and David Bohlke, published by Cambridge, as the main teaching content for the both groups. Both groups received instruction fifty minutes a week in this study, from the beginning of March to the end of May, 2013. The teaching schedule was followed by the original English teacher's syllabus. The textbooks '*Four Corners*' offers twelve topic-based units with a variety of different authentic materials. The series of *Four Corners* are an integrated four-skill course for adults and young adults who want to use English to communicate effectively in daily life. It combines proven communicative methodology with a practical outcomes-based approach. *Four Corners* features a clear presentation of vocabulary, a thorough grammar syllabus, and an everyday functional language lesson in every unit together with

systematic practice of all four skills. It places special emphasis on helping students become confident and competent speakers of English. Speaking activities are tied to clearly label measurable outcomes, enabling students to see the results of their learning and help them see their progress.

In this textbook, there are totally twelve lessons. Due to the limited time, the teacher only taught units seven to twelve in this semester. This study covered units seven to nine.

Since the purpose of this research was to improve students' vocabulary knowledge, data-driven learning approach was designed to focus on prompting near-synonyms learning. In this research, three units were used for the main teaching content. For each unit, there were at least ten pairs of near-synonyms and in total there were 34 pairs of near-synonyms, including 89 words in the experimental teaching. The topic and the pair of near-synonyms are listed in Table 3.5.

Table 3.5 *Near-synonyms List of the Study*

Unit	Topic	Paris of near-synonyms
Unit 7	New Ways of thinking	1. human, person, people 2. see, watch, look at 3. spend, take, cost 4. tour, travel, trip 5. eventful, significant, important 6. creative, imaginative 7. advantage, benefit 8. alter, change 9. across, cross 10. date, appointment 11. field, course, court, ground, ring 12. heaven, paradise
		1. continue, begin, start 2. regard, consider, think

Unit 8	Lessons in life	3. understand, comprehend, realize 4. spell, pronounce 5. diligent, industrious 6. decision, determination, resolution 7. disrupt, interrupt 8. freedom, liberty 9. grade, mark, point, score 10. huge, big, large, enormous 11. opportunity, chance 12. persuade, convince
Unit 9	Can you explain it?	1. homework, assignment 2. hurt, injure, damage 3. heal, cure 4. hospital, clinic 5. inhibit, prohibit 6. option, choice 7. replace, substitute 8. permit, allow 9. game, contest, competition, match, meet, race 10. assure, ensure

3.4.2 Teaching Activities

In this research, the participating students had forty minutes for teaching near-synonyms per week. It took about two or three weeks to learn a unit. The activities used in each unit and the timetable are presented in Table 3.5.

All the teaching activities were conducted through the data-driven learning with the corpus-based concordance lines of the teaching materials. Like the pilot study, in the quasi-experiment study, hands-on activities were designed by the researcher who put emphasis on the vocabulary strategies based on the Schmitt's (2012:207-208) taxonomy.

These activities intended to make use of the corpus to enhance participants' attention on vocabulary and accessed their vocabulary knowledge.

Table 3.6 *Teaching Activities and Timetable of the Study*

Week	Date	Unit	Activities	Length of Activity
1	3/4		Conduct pretest and background questionnaire	40 mins
			Conduct orientation	10 mins
2	3/11	Unit 7	'*human, person, people*', '*see, watch, look at*', '*spend, take, cost*' ,'*tour, travel, trip*'	50 mins
3	3/18		*'eventful, significant, important'* ,*'creative, imaginative', 'advantage, benefit*', *'alter, change'*	50 mins
4	3/25		*'across, cross'* ,*'date, appointment', 'field, court, ground, ring', 'heaven, paradise'*	50 mins
5	4/1	Unit 8	*'continue, begin, start', 'regard, consider, think' 'understand, comprehend, realize' 'spell, pronounce'*	50 mins
6	4/8		*'diligent, industrious'* ,*'disrupt, interrupt' 'decision, determination, resolution' 'freedom, liberty*'	50 mins
7	4/15		Mid-term	
8	4/22	Unit 9	*'grade, mark, point, score', 'huge, big, large, enormous', 'opportunity, chance', 'persuade, convince'*	50 mins
9	4/29		'*hurt, injure, damage'* ,*'heal, cure' 'hospital, clinic'* ,*'inhibit, prohibit' 'option, choice'*	50 mins
10	5/6		'*join, attend, participate, enter', 'mature, ripe' 'replicate, duplicate, copy', 'assure, insure*'	50 mins
11	5/13		Conduct posttest	30 mins
			learners' perception questionnaire	20 mins

The instruction procedures consisted of four main parts in each class. The following section gives a brief description of each part and their purpose.

3.4.3 Teaching Procedure

Under the notion of Anderson's Adaptive Control of Thought (ACT) Model (Takač, 2008:35), three stages "cognitive, associative and autonomous" were adapted in activity. These activities also followed the traditional 3P model— presentation, practice and production. The 3P model can be coded into three stages mentioned in Anderson's model.

Orientation

Students in the experimental group had a concordance observational orientation to ensure they have no problem analyzing the concordance. As shown in Appendix A, the orientation was divided into three parts: introduce the concept of near-synonym, corpus and concordance.

Warm up

In a natural setting, learners always encounter words in context rather than in isolation. This warm-up activity not only gave students confidence but also self-efficacy to continue their learning, interest and achievement (Takač, 2008:43). In this stage, teacher used a strategy which was "guess from the lyric context" under Schmitt's (2012) and Chandrasegaran's (1980) framework. Before listening to the music, the teacher asked students three questions for inferring words' meaning and also for asking students to guess the word meaning. After listening to the pop song, students discussed with each other and then teacher generalized students' answers. In this part, teacher helped students

create word consciousness through asking three pre-listening questions.

Vocabulary presentation

In the presentation stage, teacher demonstrated the research procedures which students followed in the proceeding activities, such as how to find the natural lexicon through dictionaries or teacher-made concordances. After illustrating the concept of concordance, the teacher used near-synonym and corpus to raise students' awareness of second language vocabulary learning. Students at this time not only learned declarative knowledge but also conceptual structures in the cognitive and conative model of Young & Perkins (1995). Students will observe near-synonyms in each group and wrote down their observations on the worksheet. Once they have questions, they could discuss with their teammates. The teacher did not give the students correct answers immediately in order to activate their active participation as suggested in Stern's synthesis model (1986). Based on the Stern's synthesis framework, in this part, teacher cultivated students "active planning strategies" and "social learning strategies (Stern, 1986:411)." According to Young and Perkins (1995:150), they define learning strategies as "specialized ways of processing information that enhance its comprehension, learning or retention." To attain so-called "specialized ways", the teacher has to give students some strategy trainings in consideration of different personal learning styles. To sum up, in the presentation stage, strategy was introduced by teacher at first and then teacher gave students some hand-on activities to make sure they understood the procedures of researching on near-synonymous words.

Vocabulary Activities

Enhancing the memory of the new learn words was the purpose of this stage. Students did hand-on activities in order to ensure they understand the concept of concordance and near-synonym. The students will become more "proceduralized" through practice Anderson's ACT framework (Takač, 2008:35).

Wrap up

The teacher generalized each pair of near-synonyms patterns and asked students questions in order to make sure the participants have the ability to apply the target words in various contexts.

3.5 Control Instruction Design

Students in the control group were taught in a regular classroom setting based on the traditional teaching approach used in Taiwan. The teaching procedures are teacher-centered lecture instead of learner-centered and can be applied to various language classrooms without modification. The teacher in the control group views language learning as a habit formation and the ultimate goal is to help students achieve the mastery of language learning. There were no group discussions and hands-on activities in the control group. The teacher focused on explanation of words' meanings, grammar features, word-for-word translation and memorization of words. To put it more specifically, the teacher used Grammar-Translation Method to teach learners and put emphasis on the explanation of grammar features and the L1 translation of the target words. The teacher directly told the differences among pairs of near-synonyms.

The teaching content and the range of vocabulary were identical with the experimental group followed by the original teacher's syllabus and learners' textbook.

The instructional period was the same as the experimental group, which lasted eight weeks, fifty minutes per week respectively. The teaching materials were compiled by the teacher before class. The teacher used five on-line dictionaries (*Merriam Webster, Longman dictionary of contemporary English, Oxford dictionary, Cambridge dictionary, Collins English dictionary*) to select the word's core meanings and put the most salient meanings into teaching materials. For instance, in 'people' senses, the researcher at first listed all meanings from the five on-line dictionaries and then chose the top and second salient meaning of the target word. If the researcher found the meaning was obscure, an example would be given after meaning. The selective procedure of the word 'people' is shown in Table 3.7 and five senses of 'people' are shown in Table 3.8. The teaching materials for the control group are displayed in Appendix C.

Table 3.7 *The Selective Procedure of 'people'*

	Meaning	Merriam Webster	Longman	Oxford	Cambridge	Collins
1.	plural : In general of consider collectively	✓	✓	✓	✓	✓
2.	[verb] (of a group of people) inhabit (a place)[to dwell in] ex: *an arid mountain region peopled by warring clans*	✓	✓	✓	✓	✓
3.	[verb] fill or be present in (a place or domain) ex: *in her imagination the flat was suddenly peopled with ghosts*	✓	✓	✓	✓	✓
4	plural: the members of a family or kinship ex: *Do your people live round here?*	✓	✓	✓		✓
	plural : the mass of a community					

5	as distinguished from a special class —often used by Communists to distinguish Communists from other people ex: *poor/rich people*	✓		✓	✓	✓

Table 3.8 *Five Senses of 'people' in Selected Dictionaries*

1	plural : In general of consider collectively
2	plural : the mass of a community as distinguished from a special class —often used by Communists to distinguish Communists from other people ex: *poor/rich people*
3	plural: the members of a family or kinship ex: *Do your people live round here?*
4	[verb] (of a group of people) inhabit (a place)[to dwell in] *ex: an arid mountain region peopled by warring clans*
5	[verb] fill or be present in (a place or domain) ex: *in her imagination the flat was suddenly peopled with ghosts*

3.6 Instruments

3.6.1 Background Questionnaire

The participants completed the background questionnaire at the beginning of the experiment. The background questionnaire (see Appendix D) was designed to obtain information about participants' English learning experience and their views toward their second language vocabulary learning experience. The first part (first to ninth items) was with regard to participants' the basic information of English learning. The questions were asked, such as how long they have been exposed to English and whether they had ever lived in English-speaking countries. The second part (tenth to nineteenth items) was related to the participants' experience on using computer to learn English. Furthermore,

their English proficiency test, such as TOEIC and GEPT score, and whether they know the corpus were also asked. A student background questionnaire was given to the participants before the instruction. Such information was also helpful for collecting factors that may influence the research outcomes.

3.6.2 Self-Designed Vocabulary Test

The main test instruments (see Appendix E) are the same for all participants, consisting of a pretest, and a posttest in an identical format. The vocabulary tests measure not only learners' receptive knowledge but also learners' productive knowledge. Using multiple-choice questions and fill-in-blank questions to measure receptive and productive knowledge of different aspects may provide a much more accurate evaluation of the relative efficacy of tasks (Webb, 2005).

3.6.3 Student Feedback Questionnaire

During pretest, students were allowed 40 minutes to complete the questions without using reference books. The test only focused on near-synonym items and their expressions. The same items were used in the later posttest. In addition to test learners' vocabulary knowledge, the researcher also designed a questionnaire (see Appendix G) which was adapted from the research (Lin, Chan, & Hsiao, 2011; Yeh, 2013) to explore learners' perceptions of learning in a corpus-based DDL approach. This questionnaire was only given to the experimental group after data-driven instruction. It combined twenty three closed statements and four open-ended questions. In the closed questionnaire, it used 4-point Likert scales, 1 for strongly disagree and 4 for strongly agree. The neutral rank scale was removed to prevent the results from ambiguous attitude. The closed questionnaire consists of three parts: (1) to investigate learners' perceptions on

data-driving learning Value ; (2) learners' perceptions toward data-driven learning activity in language learning; (3) learners' perceptions on teacher's teaching. In the open-ended questions, it tapped into students' feelings and suggestions toward this methodology. Learners' perception questionnaire, which was administered to learners in pilot study, reported a Cranach α coefficient of .829 (> 0.70). Therefore, the coefficient of reliability demonstrated the post-questionnaire was adequately reliable.

3.6.4 Semi-Structured Student Interviews

In addition to the quantitative data, in order to have deeper understanding and analysis, the group interview was held after the implementation of DDL. For participants from the experimental group, the group interview was held in a classroom after whole eight week instructions. The researcher chose three participants from low and high achievers separately from the experimental group. In the group interview, the researcher prepared seven questions (see Appendix H) for participants to answer and the interviewees were allowed to say anything they want during the interview. The whole process of the group interview was recorded. Both of the researcher and participants spoke in Chinese. Moreover, the way the two group interviews held did not like formal interviews. Instead, it was more like a sharing and chatting between the researcher and participants. Thus, under this atmosphere, the participants shared lots of opinions and thoughts with the researcher, and opinions and thoughts had become important information.

3.7 Chapter Summary

3.7.1 Quantitative Data

After the experiment, data collected from pre- and post tests were documented and analyzed by using Statistical Package for the Social Science (SPSS) version 17.0 for Windows. The .05 level of significance was adapted for data analysis in this study. In order to answer the first research question, the participants' score in the pre-tests and post-tests were analyzed using paired sample *t*-test to find out whether there were significant improvements in each group. The second research question was to be answered by the same pre- and post- tests. The data were analyzed by the independent samples *t*-test to investigate whether there were significant differences regarding data-driven instruction on high achievers' near-synonymous vocabulary knowledge. In addition to the vocabulary scores, the student's perception questionnaire was given to the experimental group and then the questionnaire was analyzed descriptively in terms of means and percentages of respondents in each item.

3.7.2 Qualitative Data

For the qualitative analysis, the semi-structure student interviews were conducted in Chinese, but the transcription of interviewees' responses was translated into English.

Research measurement and statistic methods corresponding to the research questions are shown in Table 3.9.

Table 3.9

Research *Measurement and Statistic Methods Corresponding to Research Questions*

Research question	Research measurement	Statistic Methods
(1) Does the data-driven learning (DDL) approach with manual concordances would improve better in students' near-synonyms knowledge and performance in comparison to control group taught in traditional method?	◆ Pretest and posttest (the progress of the participants within the same group before and after the experiment)	Paired sample *t*-test
(2) Are there any significant differences between high and low achievers in the experiment and the control group?	◆ Pretest and posttest (the differences of the high achievers between the two groups after the experiment)	Independent sample *t*-test
(3) What are learners' reactions and preferences regarding using the corpus to learn vocabulary?	◆ Questionnaire and semi-structure interview	Frequency analysis

國立政治大學
National Chengchi University

CHAPTER 4

RESULTS AND DISCUSSIONS

This chapter presents the research results and discussions of this study. In this study, both quantitative and qualitative data analysis were performed. In terms of methods of quantitative research, the scores of the control and the experimental groups were analyzed by two statistical methods 'paired samples *t* test and independent samples *t* test'. In order to have deeper understanding of participants' perceptions, attitudes, preference and opinions toward data-driven learning (DDL), the methods of qualitative research, questionnaire and interview, were adopted as well.

This chapter begins with the result of the pretest, followed by the results of posttests. Later, the analysis of the collected data and the discussion of the research questions of this study are displayed into three sections: (a) the influence of DDL instruction on students' near-synonymous vocabulary knowledge; (b) the influence of DDL instruction on high achievers' performance of DDL activities in each experimental group; (c) participants' feedback on DDL vocabulary instruction.

4.1 Pretest

In order to make sure that the control group and the experimental group were comparable prior to the 8-week instruction, the pre-tests were administrated to all freshmen. The pretest includes the assessment of word usages on thirty-four pairs of near-synonyms presented in sentences and a dialogue. The total items for assessment are 50, and the test includes multiple choice, matching and conversation.

Independent sample *t*-test was applied to investigate treatment effects to select classes with the same language proficiency. The means, standard deviations and between-group differences of the pretest of two chosen groups are presented in Table 4.1.

The results showed that there were no significant differences between control group and experimental group ($t = .172$, $p > .05$). The two classes were comparable in the ability of their proficiency knowledge of near-synonym; therefore, Class 1 and Class 2 were assigned to the control group and experimental group.

Table 4.1 *Independent sample T-test of pretests*

Group	Mean	N	SD	t	Sig. (2-tailed)
Class 1	48.70	43	8.40		
				.172	.768
Class 2	48.37	43	9.18		

4.2 Posttest

4.2.1 Learners' Learning Achievement Scores

4.2.1.1 Paired samples *t* test

Paired samples *t*-test was conducted to examine the treatment effect on participants' near-synonymous vocabulary knowledge for the control group and experimental group individually. The means and standard deviations for within group differences of the posttests are presented in Table 4.2 for each group. An alpha level at $p = .05$ was considered significant.

Table 4.2 *Paired sample T-test in the two groups*

Group	Test	Mean	N	SD	*df*	*t*	*Sig*. (2-tailed)
Control	Pretest	48.70	43	8.41	42	4.525	.000***
	Posttest	55.35	43	9.05			
Experimental	Pretest	48.37	43	9.18	42	8.505	.000***
	Posttest	65.44	43	11.50			

*** $p < .001$

The results showed that participants in the experimental group made significant improvement by 17.07 points in the posttest ($t = 8.505$, $p < .05$). As for the control group, participants also made significant progress in posttest ($t = 4.525$, $p < .05$), the mean score of the posttest increased by 6.65 points compared with that of the pretest. In other words, participants in both groups showed great progress in their near-synonymous vocabulary knowledge after near-synonym instruction with or without corpus-based data-driven learning. The pretest mean scores of the

experimental group were lower than that of the control group; however, the posttest mean scores of the experimental group were higher than that of the control group.

To sum up, to answer the first research question, there were significant differences in learners' posttest scores between the experimental group and the control group. The corpus-based data-driven learning group outperformed the traditional in-class teaching group in their learning achievement scores. Such a finding confirms study of (Sun &Wang, 2003; Chan & Liou, 2005; Koosha & Jafarpour, 2006; Yeh, Liou & Li, 2007; Alex, 2009; Ewa, 2011; Serkan, 2011) that data-driven learning enhances participants' second language vocabulary acquisition and data-driven learning approach performed better than traditional teaching methodology. The possible reason might be that the data-driven learning approach helped participants pay attention to the word forms and retrieve the word meanings from their memory. Moreover, another possible reason may be that data-driven learning is not only a process-oriented but a product-oriented approach. Participants made effort to investigate the target words with his or her classmates and infer the meanings and usages within contexts. Participants had a sense of participation through the process and had a strong impression on the target words after the whole process. DDL not only promotes creativity and self-discovery learning among learners, but also presents to the learners by multiple exposures within contexts (Batstone, 1995). The results also echoed the statement that "learners have a big improvement after corpus-based teaching instructions (Tribble & Jones, 1990; John, 1994; Tribble, 1996; Kettemann, 1995)."

4.2.1.2 Independent sample *t*-test

Since the treatment yielded significant effects in the posttest of data-driven learning in the experimental and the control group, their specific improvements in high and low achievers were further examined.

Independent sample *t*-test was applied to examine the treatment effects on posttest between high and low achievers in the control and the experimental groups. An alpha level set at $p = .05$ was considered significant. The means and standard deviations of the between-group high and low achievers' differences are shown in Table 4.3 and Table 4.4.

Table 4.3 *Independent sample T-test of posttest of high achievers in the two groups*

Group	Mean	SD	*df*	t	*Sig*.(2-tailed)
Control	58.40	11.44	28	2.735	.011***
Experimental	69.47	10.70			

*** $p < .001$

Table 4.4 *Independent sample T-test of posttest of low achievers in the two groups*

Group	Mean	SD	*df*	t	*Sig*.(2-tailed)
Control	54.92	11.41	28	1.748	.197
Experimental	59.53	8.39			

In the aspect of high achievers, the results showed that the differences between the control group and the experimental group were significant ($t = 2.735$, $p < .025$). The mean score of the experimental group ($M = 69.47$) outperformed that of the control group ($M = 58.40$), which indicates that when learning the near-synonyms, EFL high proficiency learners perform better with data-driven learning approach as scaffolding than high achievers who were taught by the traditional method.

This finding is not in contradiction with that of the empirical study in previous research. For high and low achievers with DDL approach, this finding confirms the study of Lin (2005), though there are important differences regarding other aspects of the study. This result lends some credence to the hypothesis that different language proficiency and different language competence might have impact on the effect of data-driven learning (e.g. Tseng, 2002). High achievers in DDL group seem to be indicative of the fact that inductive learning can lead to greater improvement to second language vocabulary learning. However, the data suggests that DDL does not only improve learners' receptive skills, but also improve their productive skills, which was not found in Lin's study.

In the aspect of low achievers, the results showed that the mean score of the experimental group (M = 59.53) was slightly higher than that of the control group (M = 54.92). However, the differences between the control group and the experimental group were non-significant (t = 1.748, p >.05), which rejected the hypothesis that when learning the near-synonyms, EFL low proficiency learners perform better with data-driven learning approach than low achievers who were taught by the traditional method.

Since using DDL to teach low achievers has been a controversial issue, the direct comparisons must be treated with caution. The present study enhances the previous studies by providing a much more detailed examination of DDL. In Hadley's (2002) study, he was assured by colleagues in Japan that any attempt to use DDL with beginners was doomed to failure even though these same colleagues had never tried to use DDL to learn vocabulary. Other than Hadley, some researchers (Cobb, 1999; Allan, 2006; Koosha and Jafarpour, 2006) also assumed that DDL is not helpful for low achievers. The result of present study conforms to the assumption. The possible

reason is that low achievers might simply not have sufficient analytical and linguistic skills to cope with the complexity of authentic data of a second or foreign language. The deeper reasons were further investigated in the qualitative results.

4.2.1.3 Summary of result

Results of the pretest and posttest comparisons revealed that both groups improved their vocabulary knowledge, after receiving near-synonymous instructions, with or without using data-driven learning approach. Although the control group made some progress in the posttest, the gains of the experimental group were significantly greater than those of the control group. According to Figure 4.1, we can see that before the data-driven instructions, the participants' entering behaviors were the same which meant they had the same vocabulary knowledge before receiving the teaching treatment. Later, after eight weeks of data-driven learning training, the participants in the experiment group acquired more vocabulary knowledge than those of the control group.

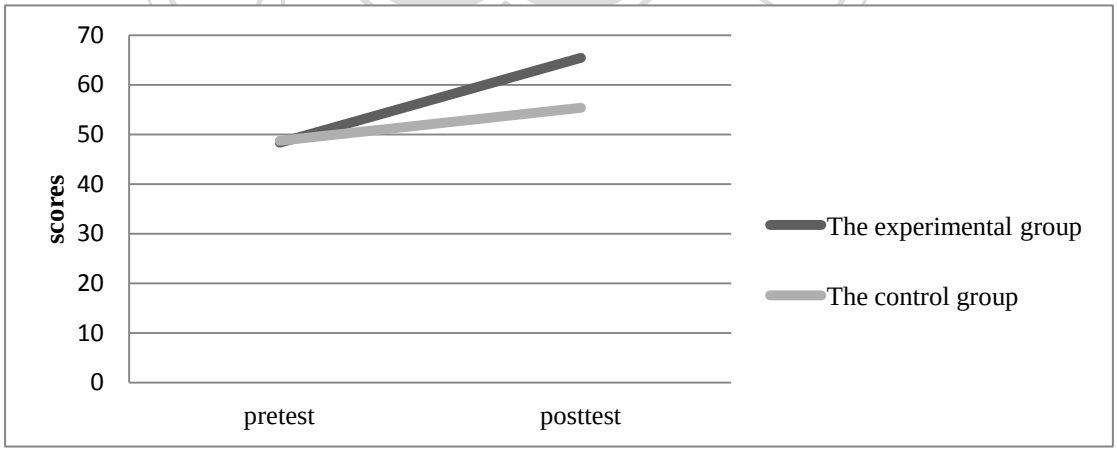

Figure 4.1 Pre- and Post-test Scores

The results of between-group comparison of both high achievers showed that high proficiency EFL learners improved their performance in vocabulary knowledge,

after being treated with or without data-driven learning approach. Data-driven learning approach and traditional teaching method both had influence on learners' vocabulary knowledge. High achievers in the experimental groups made more progress than those of the control group. Overall speaking, the participants in the experiemtal group outperformed those of the control group.

Figure 4.2 showed the progress of high achievers and low achievers in their pre-and post-tests. The high achievers' mean score in the experiment group had a lot of progess (incresed from 52.53 to 69.47) than that of the control group (incresed from 52.8 to 58.4). Similarly, the low achievers' mean score in the experiment group had a lot of improvement (increased from 45.06 to 60.12) than that of in the control group (increased from 44.62 to 54.92). The data showed that no matter who the participants were, the high or low achievers made a lot of imporvement in the experimental group rather than the control group. Additionally, when comparing the between-group progress gains, the high achievers' mean score still had more improvement (increased from 58.04 to 69.47) than low achivers' mean score (increased from 54.92 to 59.53). This result suggests that DDL might have much more influence on high achievers than low achievers.

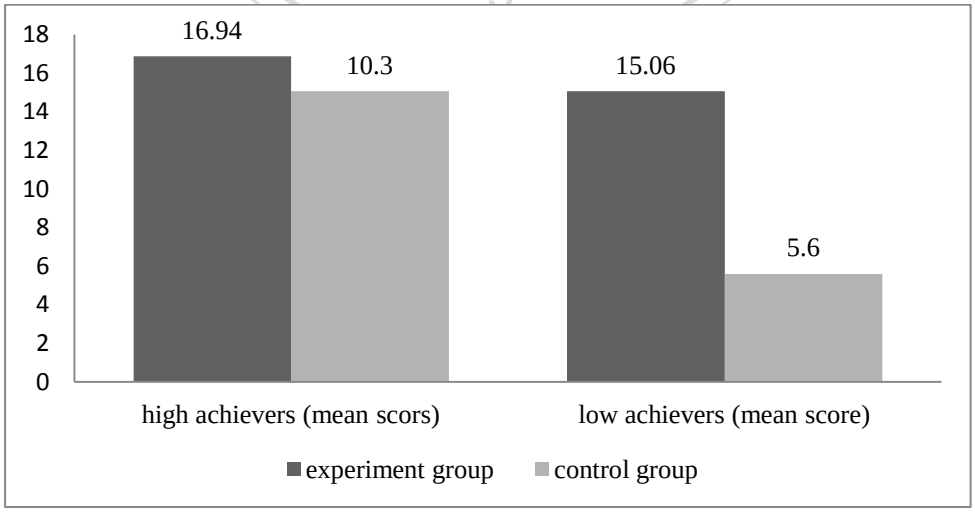

Figure 4.2 Pre- and Post-test Results of Progress Gains

Based on the results of within-group comparison or between-groups, it showed that data-driven learning is a more ideal scaffolding teaching methodology to cultivate EFL learners' knowledge of vocabulary knowledge of near-synonyms. The present study provides a valid proof in agreement with the fact that "data-driven learning enhanced participants' second language vocabulary acquisition."

4.2.2 Questionnaire results: Learners' Perceptions

The participants were surveyed with the feedback questionnaire at the end of the experimental teaching. Questionnaires were administrated for the purpose of answering the third research question. There are totally 43 participants in the experimental group. However, there were four invalid questionnaires where 10% or more of the items are left blank. There were 39 valid questionnaires returned. Their results will be discussed in the following section. The options SA 'strongly agree' and A 'agree' belongs to positive attitude, whereas SD 'strongly disagree' and D 'disagree' refers to negative attitude in this study.

Learners' attitude can be divided into three factors in this study. The first factor is 'learners' attitude toward DDL' which based on their affective and cognitive reactions. The second factor is 'what corpus-based instructional activities facilitate language learning'. The third factor wants to know how learners feel toward teachers' teaching. The data was statistically analyzed and presented in three parts to answer the third research question.

4.2.2.1 Learners' attitude: *DDL Learning Value*

Factor 1 is to measure learners' reactions on their cognitive and affective perceptions toward data-driven learning. The results of the questionnaire were shown in Table 4.5. Question 2,4,5,6,8, and 9 are concerned with the participants' cognitive

reactions; Question 1,3,7,10,11,12, and 13 are about the affective reactions.

Items 1 to 13 focused on the participants' general views towards data-driven learning approach. Most of the participants agreed that they like to learn English by using the computer technology in class (86.7%). Over seventy percent (71.8%) of students agreed that activities through computer technology can help them in English learning. 53.4% of participants felt that using corpus-based paper-printed concordances to learn vocabulary can enhance their interests. Tribble and Jones (1990) claimed that concordance lines can accelerate learners' motivation. The source of material for exercise is retrieved from real life spoken or written data rather than contrived by teachers increases motivation, "as it gives learners immediate contact with the target language in use (Tribble & Jones, 1990:36)." However, in this study, although the result in item 3 was positive, some of low achievers disagreed with the fact that data-driven approach increased their motivation. During the interview, the low achievers claimed that they did not have a lot of vocabulary knowledge so that they felt pressured when they did the activity in class. Because of lacking confidence in learning English, the majority of low achievers are passive learners who have low autonomy. The more intrinsic motivated and autonomy a student is, the more they will recognize the relevance and usage of whatever methods, materials, and approaches a teacher uses to their development as a language learner (Deci & Flaste, 1995; Allen, 2012). Hence, this result was opposed to the Tribble and Jones's (1990) research.

In addition, in this part of questionnaire of the current study, 87.1% of participants stated that they will avoid certain errors of vocabulary usage in the future after using the corpora. Concordance lines provide an important corrective to the portrayal of English books, which is sometimes inaccurate (Lewis, 1993; Mindt, 1997;

Fox, 1998) or based on writer's intuitions (Tomlinson, 1998). Cognitively speaking, the result revealed that participants can be aware of the importance of using corpus to learn English vocabulary.

Based on the data-driven approach, 53.8% of the participants thought they became more focused in class because they had to focus on observation and analysis of the data. Willis (1998) stated that once participants' attention has been drawn to the meanings, uses and functions of the target language, learners are more likely to notice and reflect on 'further occurrences of the language items that have been salient through the concordances (p.73).' Such positive feedback on data-driven learning value was similar to various researchers such as Koosha & Jafarpour (2006); Alex (2009); Ewa (2011); Serkan (2011) and local studies by Sun &Wang (2003); Chan & Liou (2005); Yeh, Liou & Li (2007).

Due to the data observation and analysis in each week, 71.8% participants felt using corpus-based paper-printed concordances to learn vocabulary can promote their language awareness and decoding ability. One possible reason is that learners who learn words from manual concordance lines not only know what the basic usages of each word, but also begin to see how naturally occurring English is utilized outside the classroom. Manual concordance lines foster critical awareness and provide an accessible way to analyze and keep track of the language information harnessed through their usage (Allen, 2012). Similarly, Willis' study (1998) claimed that data-driven learning helps learners recognize the parts playing collocation lexical phrases and realize there is more to language than just vocabulary and grammar. Kennedy's study (2003) held the same view. He pointed out that the second vocabulary acquisition process depends on the consciousness raising and learners should be exposed to authentic language materials as much as possible. Fox (1998)

also asserted that students are encouraged to develop further analytical skills by analyzing data in data-driven learning. In addition, 76.9% participants thought that data- driven learning approach enhanced their memory. It might be because they analyzed and decoded words' meanings by themselves in class. This is in complete agreement with Cobb (1999) and Boulton's (2009) results. They claimed that learners learning language through corpus-based concordance lines have better retention than through traditional methodology.

However, only 33.3% of the participants thought corpus-based paper-printed concordance was suitable for them, only 46.2% of the participants thought they can get higher grades in the test and only 33.3% of the learners felt a sense of achievement learning English vocabulary through data-driven learning approach. Thomlinson (1998) cautions instructors against assuming all the students are 'ready and willing' to learn any particular presentations the teacher demonstrate in class. Each learner is in charge of their individual classroom experience and thus the teachers "never completely control what the learner does (Tomlinson, 1998:12)." When college teachers want to use a new approach or activities in classroom, they should be mentally prepared that the learners might react negatively at first because they need time to get accustomed to the new methodology. This result is similar to Willis' (2003) assertion, he elaborated the idea that some students will never be receptive to the use of concordance lines, or any other approach to presenting the language. He tried to persuade teachers not to be disappointed when facing this situation. What the teachers teach is often not learned and what learners do often wasn't taught.

Items 11 and 12 were negative statements, in which 76.9% of the learners reported that using data-driven learning to learn English vocabulary was wasting time.

Almost half of the learners (48.7%) felt anxious when they did the vocabulary exercise in class. In participants' general feelings, only 35.9% of the learners feel satisfied with the experience of learning English vocabulary in a corpus-based learning environment. The possible reason could be that it was the first time for learners to use this approach because teachers had never used concordance in the classroom. The main problem might be that the teachers have not heard of concordance lines, so they do not know what they are, or they have heard them but do not properly understand what they are. As the same standpoint in Kilgarriff's (2009) study, he claimed that "Most English language teachers have not used them, and may not even have heard of them. This is probably true at the university level and certainly true at the high school level (p. 40)."

Overall, learners' attitude toward the data-driven learning approach showed positive. In affective aspect, they liked to use computer technology to learn English (M=2.87, SD= .73), thinking that computer technology can help them in learning English (M= 2.85, SD= .63). In cognitive aspect, through the data-driven approach, learners considered it not only enhanced their interests (M= 2.56, SD=.75), memory (M=2.95, SD=0.65) and concentration (M= 2.47, SD=0.64) in learning English, but also increased their decoding ability (M= 2.77, SD= 0.63) and help them to avoid certain errors in the future (M=3.05, SD=0.56). However, the minority of the learners thought that they can get high grades through data-driven learning (M= 2.51, SD=.68), having a sense of achievement (M= 2.26, SD=.59) when they learn English vocabulary through data-driven learning. The anxiety level was reportedly higher (M=2.51, SD=.64) in the data-driven classroom. Besides, seldom of the participants thought data-driven approach was not suitable for them (M= 2.26, SD=.60) and it wasted time (M= 2.87, SD= .8). In general, only 35% of the participants would like to

take this approach to learn English vocabulary (M= 2.64, SD= .67).

Table 4.5 *Results of Learners' attitude on Data-driven learning Value (N=39)*

Factor 1: learners' attitude toward DDL		Numabers of participant and frequency of responses				*M*	*SD*
No.	Item Description	SA	A	D	SD		
1	I like to learn English with the use of computer technology in English class.	13 (20.5%)	18 (46.2%)	8 (33.3%)	0	2.87	0.73
2	I think activities through computer technology can help me in learning English.	5 (12.8%)	23 (59.0%)	11 (28.2%)	0	2.85	0.63
3	I think using corpus-based paper-printed concordances to learn vocabulary can enhance my interest in learning English.	4 (10.3%)	16 (43.1%)	17 (43.6%)	2 (5.1%)	2.56	0.75
4	Thanks to the corpus data, I think I will avoid certain errors of vocabulary usage in the future.	7 (17.9%)	27 (69.2%)	5 (12.8%)	0	3.05	0.56
5	I become more focused in class when I learn English vocabulary through data-driven learning.	0	21 (53.8%)	15 (38.5%)	3 (7.7%)	2.47	0.64
6	Through data-driven learning approach, I think I can get higher grades in the test.	3 (7.7%)	15 (38.5%)	20 (51.3%)	1 (2.6%)	2.51	0.68
7	I have a sense of achievement when I learn English vocabulary through data-driven learning approach.	0	13 (33.3%)	23 (59.0%)	3 (7.7%)	2.26	0.59
8	I think using corpus-based paper-printed concordances to learn vocabulary promotes my language awareness and decoding ability。	3 (7.7%)	25 (64.1%)	10 (25.6%)	1 (2.6%)	2.77	0.63
9	I think vocabulary exercises through corpus-based paper-printed concordance	7 (17.9%)	23 (59.0%)	9 (23.1%)	0	2.95	0.65

	enhance my memory for English vocabulary.						
10	I think learning vocabulary with the use of corpus-based paper-printed concordance is suitable for me.	0	13 (33.3%)	23 (59.0%)	3 (7.7%)	2.26	0.60
11	I think using corpus-based paper-printed concordances to learn vocabulary in English class waste time.	7 (17.9%)	23 (59.0%)	6 (15.4%)	3 (7.7%)	2.87	0.80
12	I feel anxious when I learn vocabulary through data-driven learning approach.	2 (5.1%)	17 (43.6%)	19 (48.7%)	1 (2.6%)	2.51	0.64
13	In general, I feel satisfied with the experience of learning English vocabulary in a corpus-based learning environment. In that, I like the way when I learn English vocabulary.	2 (5.1%)	12 (30.8%)	23 (59.0%)	2 (5.1%)	2.64	0.67

Note: SA: Strongly agree. A: Agree. D: Disagree. SD: Strongly disagree

4.2.2.2 Learners' attitude: Hands-on Activities

The feedback of learning motivation, learning experience and materials was shown in Table 4.5. In Table 4.5, it is clear that learners' motivation was greatly enhanced, students understood the words and sentences in the teaching materials, and most importantly, they liked to discuss and discover the vocabulary rules and patterns with teammates.

Next, the questions in Table 4.6 will be discussed in details. Since the supporting literature of some results has been cited above, the researcher only discusses the new ones in order to avoid repeating the literature.

Table 4.6 *Results of Learners' Feedback on Language Learning Value (N=39)*

Factor 2: Language learning		Numabers of participant and frequency of responses				*M*	*SD*
No.	Item Description	SA	A	D	SD		
14	I think doing vocabulary exercises through paper-printed concordance are more interesting than doing identical exercises of traditional teaching on textbooks.	1 (2.6%)	25 (64.1%)	13 (33.3%)	0	2.69	0.52
15	I think doing vocabulary exercises through corpus-based paper-printed concordances would distract me from English words.	2 (5.1%)	22 (33.3%)	13 (56.4%)	2 (5.1%)	2.61	0.67
16	After using data-driven learning to learn English vocabulary, in the future, I would like to do other similar corpus activities in class, such as grammar learning or English writing training.	1 (2.6%)	16 (41.0%)	20 (51.3%)	2 (5.3%)	2.41	0.64
17	I can understand the vocabulary on the course material (paper-printed concordance) in English class.	2 (5.1%)	24 (61.5%)	13 (33.3%)	0	2.72	0.56
18	I can understand the sentence pattern on the course material (paper-printed concordance) in English class.	1 (2.6%)	21 (53.8%)	17 (43.6%)	0	2.59	0.55
19	In the course material, I think it is not enough to provide sufficient English examples for learners.	4 (10.3%)	28 (71.8%)	6 (15.4%)	1 (2.6%)	2.90	0.60
20	In data-driven learning activity, I feel it is great for finding vocabulary rules patterns with teammates cooperatively.	2 (5.1%)	20 (51.3%)	16 (41.0%)	1 (2.6%)	2.59	0.64
21	In data-driven learning activity, I have a sense of achievement finding vocabulary rules and patterns with	3 (7.7%)	19 (48.7%)	17 (43.6%)	0	2.64	0.63

	teammates cooperatively.						
22	In data-driven learning activity, I can learn a lot of English vocabulary rules and patterns from other peers who have high proficiency in English.	4 (10.3%)	22 (56.4%)	12 (30.8%)	1 (2.6%)	2.74	0.68
23	(a) I would like to explore corpus on computer rather than through prepared paper-based exercises.	1 (2.6%)	15 (38.5%)	18 (46.2%)	5 (12.8%)	2.30	0.73

Note: SA: Strongly agree. A: Agree. D: Disagree. SD: Strongly disagree

According to Table 4.5, 66.7% of the participants responded positively that doing vocabulary exercises through paper-printed concordance are more interesting than doing identical exercises of traditional teaching on textbooks. 38.4% learners expressed that doing vocabulary exercises through corpus-based paper-printed concordances would distract themselves from learning English words. In that, the rest of 61.5% learners thought they can focus on the English words through this activity. One explanation for this is that learners have dynamic experience when they learn vocabulary within the expansive contexts, so they are more concentrated on learning the lexical chunks entailing words, grammars, discourse as well as pragmatic conventions that shape the use of lexis (Tribble and Jones, 1990; D. Willis, 1990; Lewis, 1993; J. Willis, 1998; O'Keefe et al., 2007). However, only 43.6 % of the participants were willing to use data-driven approach in other corpus-based activities, such as grammar learning or English writing.

When asked about participants' feedback on corpus-based concordance teaching materials, most of the participants agreed that they can understand the vocabulary (66.6%) and sentence patterns (56.4%) in the course material. Furthermore, 82.1% of the participants strongly agreed that the English examples in

teaching materials are enough for them.

Items 20 to 22 dealt with participants' feedback on cooperative and collaborative learning experience. In item 20, there were more than half of the participants (56.4%) agreed that they felt great when they found vocabulary rules and patterns with teammates cooperatively. Meanwhile, in item 21 there were also more than half of the participants (56.4%) agreed that they have a sense of achievement finding vocabulary rules and patterns with peers. Compared to the item 7, only 33.3% of the participants felt that they have a sense of achievement when they did corpus-based activity by themselves. In Item 22, over 66% of the participants believed that they learned a lot of English words and patterns from high achievers. The results indicated two important findings from the cooperative and collaborative learning. One is that participants felt they not only have a great feeling, but also a sense of achievement when learning English words with teammates. The other is that participants felt they could learn vocabulary knowledge from other high achievers classmates. According to Dörnyei and Murphey (2003), they listed 'learning about each other' as one of the elements which could motivate students' learning. They also point out that 'learning about each other' is the crucial element for improving the relationship among group member. In the same study (2003), they claimed that cooperation toward common goals is the most effective method to motivate students. Haller, Gallagher, Weldon and Felder (2000) also proposed cooperative learning has positive effects on students' interdependence, responsibility and accountability. Through cooperative learning, the participants not only learned from each other, but also learned knowledge from each other.

Next, in item 23, approximately 40% (41.1%) of the participants agreed that they would like to explore corpus on computer rather than through prepared

paper-based exercises. The results showed some important findings which described in the following. First, the participants thought it is much fun when they search words through on-line corpus by themselves. Second, some participants who want to explore corpus on computer felt much more sense of participation and self-efficacy than doing paper-based exercises. Some thought it is convenient for them to use computer to collect information. Moreover, they believed that they could enhance their memory by doing actual manipulation. Third, some of them considered that examples and sentence patterns are limited in paper-based exercise because of teachers' selection. The findings are in line with Supatronant's (2005) study. Although the learning process on computer is complicated, the students felt satisfied with the end-result. However, the majority of the participants (59%) did not like to explore corpus on computer. Instead, they liked to learn English vocabulary from prepared paper-based exercises. They thought it is easy and time-saving for them to use paper-based exercises to learn English vocabulary and they can take notes on worksheets. Second, some participants liked to discuss with classmates through the paper-based exercise when they learn the words. Third, the minority of participants pointed out that when they use the computer to learn English words, they might be distracted by the Internet. Fourthly, compared to participants who would like to explore corpus on computer, participants thought they felt a sense of security because of the limited vocabulary range. In paper-based exercises, the lists of words were selected by the teacher in advance. Additionally, participants expressed that they were accustomed to learning vocabulary through paper-printed materials. They felt it was ineffective to their study if the teaching material was replaced with the on-line corpus.

In sum, learning English vocabulary through paper-printed concordance was effective. The majority of participants perceived that doing vocabulary exercises

through paper-printed concordance are more interesting than doing identical exercises of traditional teaching on textbooks (M= 2.69, SD= .52). They felt fewer distractions when they learn vocabulary through paper-printed concordance (M= 2.61, SD= .67). However, slightly more learners considered that they would like to do other similar corpus activities in class, such as grammar learning or English writing training (M= 2.41, SD=.64). Few participants would like to explore corpus on computer (M= 2.30, SD= .73). When asked the course material, over half of the participants expressed that they can understand not only vocabulary (M= 2.72, SD= .56) but also sentence patterns (M= 2.59, SD= .55). In addition, a lot of participants disagreed the statement 'English examples are insufficient to them on the course material (M= 2.90, SD= .60).' They thought the examples in course material were enough for them to find the vocabulary patterns. To the collaborative and cooperative learning aspect, participants felt the atmosphere was great (M= 2.59, SD= .64). They had a sense of achievement (M= 2.63, SD= .63) when they studied vocabulary with teammates. Through group discussion, they claimed that they can learn more vocabulary knowledge from high achievers (M= 2.74, SD= .68).

4.2.2.3 Learners' attitude: *Teacher's teaching value*

In order to control the variable of teacher's teaching attitude, the researcher designed three questions based on the third factor, and the questions are related to the aspect of the participants' perceptions toward teacher's teaching skill, teaching attitude and teaching materials. The results of students' response were shown in Table 4.7.

Table 4.7 *Results of Learners' Feedback on Teacher's teaching Value (N=39)*

Factor 3 : Teacher's teaching value		Numabers of participant and frequency of responses				*M*	*SD*
No.	Item Description	SA	A	D	SD		
27	I think I learn a lot of vocabulary knowledge from teacher's teaching.	12 (30.8%)	26 (66.7%)	1 (2.6%)	0	3.28	0.51
28	I think the teacher with an enthusiastic attitude in teaching.	20 (51.3%)	15 (38.5%)	4 (10.3%)	0	3.41	0.68
29	I think the teacher is carefully well-prepared teaching materials before class.	25 (64.1%)	14 (35.9%)	0	0	3.64	0.49

Note: SA: Strongly agree. A: Agree. D: Disagree. SD: Strongly disagree

Table 4.6 shows over ninety-five percent (97.5%) of the participants agreed they learn a lot of vocabulary knowledge from teacher's teaching. There were 89.8% of the participants felt the teacher was enthusiastic when he taught near-synonyms in class. Moreover, all of them (100%) thought the teacher was well-prepared before class. To conclude, participants felt positive toward teacher's teaching in three aspects: teaching skill (M= 3.28, SD= .52), teaching attitude (M= 3.42, SD= .68), teaching materials (M= 3.64, SD= .49).

4.2.3 Qualitative Findings: Interview

When the teacher uses DDL to teach students, they may find it hard to design the teaching materials for high and low achievers in a heterogeneously grouped class. In Taiwan EFL context, heterogeneous grouping is a common way for school to arrange students in every class. Heterogeneous grouping refers to a way of grouping students with mixed achievement levels in a classroom setting. Hence, every student in Taiwan's classrooms has different learning styles and skill levels. While some high

achievers can speed-read through their text book, others may take longer to comprehend the teaching materials. When teaching mixed level classes using a flexible teaching approach such as data-driven learning approach, teachers may address the learning needs of their students with different English proficiency using various data-driven learning activities. By identifying both low and high achievers' perceptions of DDL activities, teachers can desigh activities that not only accommodate their readiness level but also attract them to participate. Hence, different English proficiency students can benefit from the data-driven-based activities. The researcher interviewed three high achievers and three low achievers in this study. The following section provides the information from the interviews with high and low achievers about their perception toward data-driven learning approach.

4.2.3.1 High Achievers: prefer challenging tasks

When interviewing with the high achievers about their learning experience of using data-driven approach to learn English near-synonyms, most of them were glad that they had the opportunities to learn a new methodology in this course. Through the vocabulary pretest in this study, the high achievers showed that they had a larger vocabulary size than the low achievers. Before learning data-driven learning, they had learned the near-synonyms when they were senior high school students. They pointed out that they merely learned near-synonyms through the translations. For example, the near-synonyms 'devote' and 'dedicate', the teacher directly elaborated the differences between those two words. This teaching approach is lacking of inductive thinking so that students forget the words. After eight weeks instructions, the high-achievers indicated that they felt challenged in using data-driven learning to learn near-synonyms at first. They felt anxious because they were not fammilar with the

new teaching approach and they wanted a clear and definite answer. They did not like ambiguous answers. One of the high achevers said : “ At first, I cannot tell the differences between the two-near-synonyms because some words are diffcult to me when I looked into the meaning in the context. After eight weeks, I felt that I can distinguish a little bit but I still needed teacher’s help to confirm my answer. However, after teacher’s explanation, for me, I am confused with some words because I think the usages and meanings are the same. I need a correct answer.” In order to answer the high achiever’s question, teachers need to empahsize that the meanings of near-synonyms are not two sides of the same coin. The meanings have gradable scale. For example, the pairs of *take part in, enter*, and *join*, have a gradable differneces among those words, from the most formal to the least formal. Another high achievers continued contending that “ Althogh eight weeks seems like a very long time, it is actually a short period. It takes more than 8 weeks to get used to the data-driven learning.” One suggested the instuctor to lengthen the sentences on course materials so that they could see more contexts within the tagart word.

When asked the comparison between the traditional learning methology and data-driven leanring, all high-achevers had a positve attitude toward data-driven leanring. They thought this method could not only help them learn Englsh by themselves when they were engaging in self-learning, but also improved their English proficiency. But they won’t let the traditional approach go. They liked the balanced approach which means both of data-driven learning and traditional deductive learning are emphasized. Using the data-driven learning to learn first and then concluded lessons by deductive learning. They enjoyed discussing near-synonyms with their group members, exchanging the information and sharing their knowledge with teammates. Data-driven activities required participants to discuss the words’ meanings

and usages in a new way and also determined the comparison and contrast of pairs of near-synonyms presented to the class. This activity seemed to satisfy high achievers' need with higher level learning objective. Three high achievers expressed that they enjoyed learning English vocabulary through data-driven learning because they liked to challenge themsleves. They hoped that one day this method could help them get high scores on TOEIC or GEPT test. Moreover, one of high achievers said "When I learn some words that I feel interested in like *date* and *appointment*, I can remember them easily, having a long retention." At the same time, in the group discussion, they mentioned that some of the low achievers would not like to take part in discussing because of the lack of confidence and basic vocabulary knowledge. One of the high achievers said: "low achievers didn't want to join the group discussion until the teacher changed his way of presentation. Using ramdomly selection mechanism is a good way to force them (low achievers) to pay attention to the discussion. I think this activity is diffcult for them ."

In short, high achievers in this study enjoyed the corpus-based data-driven activities that provided them with challenges as well as opportunity to expand their vocabulary knowledge. Because of the lack of confidence, they tended to use data-driven learning first and then traditional deductive learning as a supplemtary method to learn English vocabulary. They also needed to look up in the dictionary in the long run to make sure the answers are right.

4.2.3.2 Low Achievers: prefer doing tasks with Chinese translations

Low achievers ususally have more difficulties in dealing with English learning process than the high achievers. They also are more passive toward learning. However, when interviewing the low achievers of this study, most of them revealed that they recognized the value of data-driven learning approach in learning English

near-synonyms. They enjoyed the group discussion and expressed that this activities helped them learn a lot of vocabulary knowledge from high achievers because they had had few chances to discuss with peers in class. They found data-driven enhanced their memory because they spent a lot of time on observating and anlyzing the course materials. However, when it was their turn to present in class, they experienced difficulties. They felt anxious when coming to the board to perform group's task. They were afraid of mispronunciation and making grammar errors. Althogh their peers were shouting out answers to them, they still felt pressured that they could not complete the task quicker than others. They found that decoding the meaning from the given contexts was the most diffucult one when they had no Chinese translations to refer to. One of the members from his group said that "when I can see the Chinese traslations and I can use the dictionary in class, I feel much confident in performing the task."

Although data-driven learning was difficult for them, some of them claimed that they are willing to attempt to use this method to learn vocabulary. One of the low achievers stated that eight weeks' training was not enough, it needed a long time to cultivate his langauge obesrvational ability and language awareness. Moreover, low achievers remarked that they had never used this approach to learn English vocabulary. They were novice learners.They felt this approach was difficult for them because they were not good at English, lacking in basic vocabulary knowledge.The problem of lacking of sufficient vocabulary made low achievers perform poorly in this activitiy. One of the low achievers said that " I think English is not important for me. I have low proficiency in English. It dosen't matter because I can choose the job which is not related to English. For me, my learning target is to pass the course. I don't need to take the TOEIC or GEPT test." From the interview, the researcher understood the low

achievers have low intrinsic motivation. Some of them were self-abandoned in learning English. Low achievers found that using data-driven learning was not suitable for learning within whole English language textbooks. They thought it should be used in the textbooks which attach Chinese tranalstion as the supplementary. After eight weeks' training, like high achievers, all of the low achievers admitted that they had improved their vocabulary knowledge.

To sum up, through the feedback from the low achievers, it found that low achievers were also attracted to the data-driven learning that involved high cognition. However, without sufficient vocabulary and high motivation, most of them had difficulities in participating in activities that required them to decode the word meanings and usages from the given contexts. They felt confidient in learning English when they were in the group discussion. They still need Chinese translation to lower their anxiety.

After examing the feedbacks from the high and low achievers about the perceptions of data-driven learning, we can see their perceptions differ significantly. In terms of cognitive process, the low achievers preferred to engage in less challenging tasks such as passive vocabulary knowledge receiving and words' meanings memorizing while high achievers perferred the challenging ones such as analyzing words' meanings and observing sentence patterns actively. In terms of vocabulary learning, low achievers preferred activities that should involve Chinese translation as the supplementary note while as for the high achievers, they were looking for the teacher's explanation of words' usages and functions in the end.

4.2.3.3 Summary of Interview

After analyzing the results of the vocabulary pretest and posttest, it showed that the data-driven learning approach helped participants pay attention to the word forms and retrieve the word meanings from their memory. Using corpus-based teaching materials was easier to make impression of the word meanings and usages rather than relying on L1 translations. The findings presented in this chapter showed that participants have benefited from data-driven learning instruction in near-synonymous knowledge. Both high achievers and low achievers improved their vocabulary knowledge. In addition, over 70% of the participants thought data-driven learning enhanced their memory and decoding ability. However, only 36% of the participants liked to intergrate of data-driven learning into their vocabulary lessons. Last but not least, the valuable feedback on the perceptions of data-driven learning activities from both the high and low achievers has provided the instructor a new insight. When the teacher designs activities, the teacher's activities can involve different levels of course materials in order to meet the needs of different English proficiency students.

國立政治大學
National Chengchi University

CHAPTER 5

CONCLUSIONS AND IMPLICATIONS

This chapter is the summary of the main findings in the present study and is divided into four sections. First, the summary of the present study are presented. Next, the major findings of the present study are presented. Third, the limitations of this study are addressed and finally, based on the findings of the study, some pedagogical implications and suggestions for future studies are proposed.

5.1 Summary of the study

In this study, three research questions were addressed, including (1) Does the data-driven learning (DDL) approach with manual concordances improve students' near-synonyms knowledge and performance? (2) Are there any significant differences between high achievers in the experiment group and high achievers in the control group? (3) What are learners' reactions and preferences regarding using the corpus to learn vocabulary? The major findings are presented in the following sections.

5.2 The Major Findings

5.2.1 Data-driven learning approach improved participants' vocabulary knowledge

The vocabulary test results in this study indicated that participants' vocabulary knowledge on near-synonymous words achieved significant improvement after the data-driven learning instruction. Both high and low achievers benefited from the data-driven learning instruction and improved their skills in decoding words' meanings as well as in observing word usages from the contexts.

When learning vocabulary in class, data-driven learning approach helped participants pay attention to the word forms and retrieve the word meanings from their memory. Furthermore, corpus-based teaching materials provided participants with sufficient examples and contexts to train their inductive and decoding ability. After receiving eight weeks of instruction, participants had good awareness of language. Last but not least, participants liked to discuss the examples in group. When they were in group discussion, they felt self-confident and the slow learners can acquire a lot of knowledge from the high achievers. The present study found significant benefits of data-driven learning approach in learning English near-synonyms for university students. It is therefore safe to conclude that data-driven learning provides an alternative way for English vocabulary learning.

5.2.2 Data-driven learning shows different impact on participants' proficiency

The results showed that participant's proficiency seems to influence their willingness to participate in activities. High achievers with large vocabulary size were able to participate in activities that could facilitate them with higher order thinking and learning. Moreover, due to their high motivation, high achievers had patience to deal with difficult contexts. They also had enough vocabulary to use them effectively and

productively in different contexts. High achievers preferred learning by using the inductive way in the beginning to deductive way. In the end, they preferred knowing the answer by the teacher through the deductive way.

Low achievers at first did not want to participate in activity because their lacking of vocabulary knowledge limited them to the retrieval stage of vocabulary. Later, when they were given opportunity to seek help from peers, they become more confident within the group. Low achievers were also attracted to interesting pairs of near-synonyms such as *date*, *appointment* and *across*, *cross*, and liked to participate in group discussion. For some low achievers, they thought data-driven learning approach was not efficient to their vocabulary learning because it took a lot of time in learning one pair of near-synonyms. They liked to use traditional way to learn English vocabulary because they can learn more words in one class. Moreover, although the teacher in class told the students not to rely on the dictionaries, most of low achievers look up the electronic dictionaries secretly. Hence, during the interview, some low achievers suggested that the teacher could put some Chinese translations next to the course materials, especially on difficult and specialized words. The low achievers preferred data-driven learning activities that provided them opportunities to discuss with peers and a new approach for learning English vocabulary.

5.2.3 Students' Positive Reactions

Most of the participants had positive feedback on data-driven learning and found data-driven learning not only helped them concentrate on their learning, but also enhanced their memory, decoding and language ability. During the eight weeks of instruction, they enjoyed discussing with classmates in their group. When they discussed with teammates, they felt confident in finding the words' patterns and had a sense of achievement. The data-driven learning approach was seen by the participants as one of the contributors to scaffold their learning alongside the teacher or a more competent peers.

From the perception questionnaire, it was also found that some and low proficiency participants who had low motivation were unwilling to participate in data-driven learning activities. This phenomenon was from the influence of peers, not the instructors or the learning materials. Some participants stated that they felt anxious when they did the activity by themselves because they did not know how to perform the activity. To the researcher's surprise, the majority of the participants did not think they could get higher grade by doing the data-driven learning and most felt data-driven learning was not suitable for them. In addition, they were reluctant to do another activity through the data-driven learning. They thought data-driven learning was difficult and time-consuming. However, most of the participants admitted the value of data-driven learning. They felt data-driven learning not only helped them avoid certain errors of vocabulary usages in the future, but also enhanced their motivation in learning English vocabulary. The loud and diverse suggestions given by the whole class often limited the time and imposed them with pressure. It was suggested by some high and low achievers that instead of putting four pairs of near-synonyms in one class, the course materials should be limited to one or two pairs only. They thought the contexts on course material were difficult for them as well.

5.3 Pedagogical implications

The present findings contribute to the field's understanding of the forces acting on using data-driven learning to learn near-synonyms. Based on the findings, two pedagogical implications for EFL corpus-based near-synonyms vocabulary teaching at college are proposed. It is hoped that the results of this research demonstrate an example of how to incorporate data-driven learning into English near-synonyms instructions and expand students' vocabulary knowledge in English learning in Taiwan EFL context.

5.3.1 Adaptive Data-driven Learning Approach in English Near-synonyms learning

According to the results, the feedback of the high and low achievers on their perceptions of data-driven learning activities showed that different proficiency students possessed different thinking skills to master their vocabulary learning. Unlike traditional teaching which consists of lectures passively received by the students, data-driven learning offered an opportunity to make students participate actively in vocabulary learning through doing corpus-based hands-on activities. For example, instead of being told about the word meanings directly in the beginning, students could apply their word knowledge in different contexts in the data-driven learning induced by themselves. Hence, teachers should take different proficiency into consideration to arrange data-driven learning vocabulary activities for better vocabulary learning.

5.3.2 Instructional Activities and Design of Lessons

Instructional activities and design of lessons are two important elements in vocabulary teaching and learning. The key to successful vocabulary learning is how the teaching activities are carried out and how the lesson plans are manipulated. If the instructor knows the learning features well, then the system can become more effective for learners. The followings are some instructional activities and lesson designs in order to provide effective learning environment:

1. The findings of this study showed that most of participants, especially low achiever were attracted to interesting examples and word meanings of near-synonyms, such as *cross, across.* They seemed to participate more in the activities and remembered the words. Therefore, when design near-synonyms vocabulary activities, it is suggested to integrate the fun aspects of the targets words such as showing interesting examples or

some authentic news happened in Taiwan to engage students to participate. For example, in the present study, the teacher taught the words *huge, big, large,* and *enormous* in class. After the students induced the answers by themselves, the teacher explained the words' differences and concluded

by the authentic materials which were extracted from Apple Daily News English Column (figure 5.1). The teacher used Apple Daily News English Column as a supplementary material to help the students understand that near-synonyms could be found not only in the textbook but also in some other reading materials such as magazines or newspapers.

Q: big、giant、huge、great、immense、massive

2011年02月06日 讚 0 +1 0 更多專欄文章

都有「大」、「巨大」的意思，請問在使用上有何差別？而後面所形容的事物有特別指定嗎？

線上收聽請按

※A: big、large 意義和用法相同，都在形容尺寸、數字、數量、範圍、力量、程度……比一般水準「大」，只是 large稍微比 big 正式一點點。

●Sally has blonde hair and big blue eyes.
莎莉有一頭金髮和一雙藍藍的大眼睛。

●The billionaire lives in a large mansion.
那個富翁住在一棟大豪宅裡。

※huge、giant 形容「非常大」，也就是比 big、large 感覺再大些。great 表示「大到超乎尋常」。

●The director's last film was a huge success.
那位導演的最新電影票房極佳。

●The village was destroyed by a giant wave.
那個村莊被一個大海嘯摧毀。

●A great crowd gathered in the square.
大批群眾聚集在廣場上。

※massive 用來形容「大塊」、「粗重」的程度。immense 則是「無邊無際的，廣大的」。

●The bus was crushed by a massive boulder.
那輛巴士被一顆巨岩砸毀。

●The police are under immense pressure to find the killer.
警方受到無比的壓力要找出兇手。

Figure 5.1 Apple Daily English Column

2. Dividing students into groups is an important technique when teachers want to conduct data-driven learning approach in their classrooms. Traditionally speaking, teachers may ask the whole class to answer in union, it is easy to carry out such activity in a traditional deductive classroom than in a data-driven learning class. The students are passive learners. However, data-driven learning puts emphasis on students' active participation. Hence, students are divided in groups so that every learner is given chances to practice decoding ability and cultivating language awareness. In addition, according to this research, in the group discussion, the results showed that low achievers learn vocabulary knowledge from high achievers. Before dividing students into several groups, the teacher has to carefully design the group members. Each group should include students with various levels of proficiency. For example, the instructor in present study divided learners into eight groups, six learners in one group. Each mix-level group has its own group leader who assigns the work to each member.

3. Based on the concept of zone proximal development (ZPD), which proposed by Vygotsky (1963), learners are encouraged to be immerged in the environment, which is full of materials that bridge the gap between what they know and what they know after trying (Chang, 2013). The challenge could be fulfilled by providing learners with various scaffolding materials when learning. In the study, the course materials, corpus-based concordances, have positive effect on EFL learners' learning of the near-synonyms. Therefore, in the process of concordance selection and course material compilation, teachers have to choose concordance lines carefully. They have to conform to the principle "i+1" and try to avoid picking up ambiguous sentences from the corpus before class. Doing research and compiling course materials are the trickiest part for teachers who want to use data-driven approach to teach near-synonym in class.

When doing concordance lines selection and edition, teachers are recommended to select the most frequent words, collocations and the marked usages which are based on the salience by corpus tool such as *Sketch Engine* or *Word Smith*. Moreover, teachers are also suggested to think thoroughly possible words based on EFL learners' prior knowledge to decide what kind of near-synonyms are more appropriate for EFL college students' needs, which bridge the gap between what they are learning and what they are trying to pick up by applying what they have known. For example, in this study, when the instructor selected near-synonyms and designed the course materials, the instructor took the two versions of senior high school English textbooks in Taiwan into reference. Last but not least, the course materials have to be arranged by a gradable sequence from easy, simple words or sentences to difficult or tough ones. Because learners are novice at data-driven learning approach, they must spend few times getting with new learning methodology.

4. Good classroom management is conducive to manipulating data-driven learning in class. Data-driven approach needs a large amount of time in observing and discussing the examples. In the process of discussing, learners are easily distracted and chat with their classmates. How to keep good order in class is also the crucial element in this approach. The researcher suggests that teachers can introduce a bonus system like giving credits to enhance learners' attention even though they are college students.

5.4 Limitations and Delimitations

The present study investigates the effects of using corpus-based data-driven learning to teach English near-synonyms at a college level. The process of conducting the experiment reveals some limitations the researcher encountered. There are six limitations. First of all, due to the time limitation, participants are exposed to DDL for only 8 weeks.

Second, each class lasts only fifty minutes. When participants need more time to come up with their answers, it often slows down the lesson. The restricted access time to the DDL also limited instructor to provide extra instruction time to practice with DDL activities. Furthermore, because the researcher is not the original teacher in English classroom, the researcher's teaching schedule is limited, having to compress of time in order to keep regular class going smoothly with a fixed timetable. Third, teacher as the researcher might cause bias. In this study, the teacher teaches both groups and then analyze data in a subjective way. The situation may leave the researcher/teacher non value-neutral. This bias is somewhat eliminated by the use of questionnaire to measure teachers' teaching value in this study. The questionnaire results showed that all participants have positive attitude on teachers' teaching value. Fourth, the range of vocabulary used in this study was mostly limited in the textbook. The words are taken from learners' textbooks. Therefore, the test results cannot be generalized to verbs, adjective, noun and adverbs. Also, the design of the vocabulary test was to test participants' near-synonymous vocabulary knowledge. Thus, the findings of this study cannot be applied to general vocabulary achievement test. Fifth, high achievers in the experimental group will make more progress than that of the control group. The reason might be the high achievers know most of the target words prior to the instruction; the ceiling effect leaves them with little room for improvement to reach the significant level in the posttests. Besides, some of the target words pre-selected by the teacher in the teaching materials might already exist in high or low achievers' prior knowledge. This situation also leads to ceiling effect. Last but not least, the instructor has to understand and explore all the near-synonyms before the class because sometimes the instructor might feel confused among some near-synonyms. It is suggested that when the instructor compile the corpus-based course material, the instructor can work with colleagues to lessen the burden.

5.5 Suggestions for future study

The present study intends to investigate whether corpus-based data-driven learning instruction has effects on learners' near-synonymous vocabulary knowledge. The results from the present study suggested participants in data-driven learning group outperformed those who participated in the traditional class in the posttests. There are six recommendations for future study. First, the subjects of this study can be expanded. There were only 88 students in this study. The sample size of these 88 students is not big enough to represent all the freshman students. More participants or controlled group can be included in future investigation. In addition, this research can be used to conduct in different colleges, such as the comparison between college of science and commerce. Second, for the equality of teaching time, learners in both groups receive forty minutes to practice near-synonyms. However, the time allotted for implementing data-driven learning in vocabulary learning is too limited that the learners do not have enough time to observe concordance and discuss with their teammates. For future investigation of the application of data-driven learning in near-synonymous vocabulary learning, teaching time should be extended so that learners can get full practice through discussion and observation. Third, the present study are more focused on students' receptive vocabulary; it is highly recommended that future studies which further explore the impacts of data-driven learning on near-synonyms applied to teaching of speaking and writing. In addition, more innovative designs of the activities or more formats of the course materials can be explored further. Fourth, the target words selected are confined to words that mentioned in the textbook in order to keep the school timetable smoothly. Therefore, for full exploration of word learning, other near-synonyms which are the most confusing words for learners are recommended to be included as target words. Lastly, this study adapted an experimental research methodology that it discusses the inference between

dependent variable and independent variable. The qualitative data, with only questionnaires and interview, are not sufficient. The participants' learning achievement might be influenced by other factors or variables, such as motivation, learning style and anxiety. Therefore, further research is suggested to explore the case study or action research to gain a more comprehensive view of the effects of data-driven learning on participants' vocabulary learning.

REFERENCES

Alex, G. (2009). Using online corpora to develop students' writing skills. *ELT Journal, 63*(4), 363-372.

Allan, R. (2009). Data-driven learning and vocabulary: investigating the use of concordances with advanced learners of English. *English Language Teaching Journal, 63*(1), 23-32.

Allen, B. (2012). Concordance lines without no fear. *Proceedings of the 19th international symposium on English Teaching,* 85-96. Taipei: Crane Publishing Company Limited.

Batstone, R. (1995). Product and process: Grammar in the second language classroom. In M. Bygate, A. Tonkyn, & E. William (Eds.), *Grammar and the language teacher* (pp. 224—236). London: Prentice Hall.

Boulton, A. (2008). DDL: Reaching the parts other teaching can't reach? In A. Frankenberg-Garcia (Ed.), *Proceedings of the 8th Teaching and Language Corpora Conference* (pp. 38—44). Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisboa.

Boulton, A. (2009). Data-driven learning: reasonable fears and rationale reassurance. *Indian Journal of Applied Linguistics, 35*(1), 81-106.

Boulton, A. (2010). Data-driven learning: taking the computer out of equation. *Language Learning, 60*(3), 534-572.

Chandrasegaran, A. (1980). Teaching the context clue approach to meaning. *Guidelines, 3,* 61-68.

Chan, T. P., & Liou, H. C. (2005). Effects of CALL Approach on EFL College Students' Learning of Verb-Noun Collocations. *Computer-Assisted Language Learning*, *18*(3), 231-250.

Chang, M. L. (2013). *The effects of decodable text instruction on word learning of fourth-grade EFL learners*. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.

Chapelle, C. A. (2001). ELT, technology and change. In A. Pulverness (Ed.), *IATEFL: 2001 Brighton conference selections* (PP. 9—18). Whitstable: IATEFL

Chen, H. J. (2004). *Developing an English Collocation Retrieval Web Site for ESL Learners*. The review of existing collocation teaching and learning resources, 25-34. National Taiwan Normal University, Taipei, Taiwan.

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

Chief, L. C., Huang, C. R., Chen, K. J., Tsai, M. C., & Chang, L. L. (2000).What can near synonyms tell us? *Computational Linguistics and Chinese Language Processing* (1), 47-60.

Ciezielska-Ciupek, M. (2001). Teaching with the Internet and corpus materials: Preparation of the ELT materials using the Internet and corpus resources. In B. Lewandowska-Tomaszczyk (Ed.), *PALC 2001: Practical Applications in Language Corpora* (pp. 521—531). Frankfurt: Peter Lang.

Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hand-on concordancing. *Computer Assisted Language Learning*, *12* (4), 345-360.

Cobb, T. (2003). Do Corpus-based Electronic Dictionaries Replace Concordances? In B. Morrison, G, Green., & G. Motteram (Eds.), *Directions in CALL: Experience, experiments, evaluation* (pp. 179—206). Polytechnic University: Hong Kong.

Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21th? *TESOL Quarterly, 34* (3), 548-560.

Crail, F.I.M., & Lockart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Cummins, J. (2000). Academic language learning, transformative pedagogy, and information technology: Toward a critical balance. *TESOL Quarterly*, *34*(3), 537-547.

Curse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

Deci, E. L., & Flaste, R. (1995). *Why we do what we do*. New York: Penguin.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature language system. *Studies in Second Language Acquisition, 17*(3), 379-410.

Divjak, D. (2006).Ways of intending: delineating and structuring near-synonyms. In S. Th. Grice., & A. Stefanowitsch (Eds.), *Corpora in cognitive linguistic: corpus-based approaches to syntax and lexis* (pp. 19—56). New York : Mouton de Gruyter.

Divjak , D., & Gries, S. (2006). Ways of trying in Russian: clustering behavioral profiles. *Corpus linguistics and linguistics theory, 2*(1), 23-60.

Dörnyei, Z., & Murphy, T. (2003). *Group dynamics in the language classroom*.

Cambridge: Cambridge University Press.

Doughty, C., & Varella, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85—114).Cambridge: Cambridge University Press.

Edmonds, P., & Hirst, G. (2002). Near synonyms and lexical choice. *Computational linguistics, 28*(2), 105-144.

Ellis, R. (1991). *Second Language Acquisition and Second Language Pedagogy*. Avon: Multilingual Matters.

Ewa, D. J. (2011). The use of language corpora and concordancing software to improve grammatical competence in teaching English as a foreign language. *Sino-US English Teaching, 8*(12), 754-765.

Fox, G. (1998). Using corpus data in the classroom. In B. Tomlinson (Ed.), *Material development in language teaching* (pp. 25—43). Cambridge: Cambridge University Press.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*(3), 301-319.

Hadley, G. (2004). Sensing the winds of change: An introduction to data-driven learning. Retrieved from April 24, 2013, from http://www.nuis.ac.jp/~hadley/publication/ /windofchange/windsofchange.htm.

Haller, C. R., Gallagher, V. J., Weldon, T. L., Weldon, T. L., & Felder, R. M. (2000). Dynamics of peer education in cooperative learning workshops. *Journal of Engineering Education, 89*(3), 285-293.

Halliday, M. A. K. (1976). *Lexis as a linguistic level*. In C. Bazell, J. Catford, M. A. K., Halliday, & R. Robins (Eds.): in memory of J. R. Firth. London: Longman.

Higa, M. (1963). Interference effects of intralist word relationships in verbal learning. *Journal of Verbal Learning and Verbal Behavior*, *2*(2), 170–175.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp. 47—69). Oxford: Oxford University Press.

Hsieh, W. L. (2008). *"See", "Watch" and "Look at": Teaching Taiwanese EFL students on a corpora-based approach*. Unpublished master's thesis, National Chengchi University, Taipei, Taiwan.

Huang, C. R, Chung, S. F, & Su, I. L. (2008). Durative event: a comparison of 趕

gan3 and 搶 *qiang3*. *Proceedings of the Chinese lexical semantic workshop 2008* (CLSW), Singapore, 42-50.

Huston, S., & Francis, G. (1998). Verbs observed: A corpus-driven pedagogical grammar. *Apply Linguistics, 19*(1), 45-72.

Huston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Johns, T. F. (1986). Micro-concord: A language learners' research tool. *System, 14*(2), 151-162.

Johns, T. F. (1991). Should you be persuaded: Two examples of data-driven learning. *English Language Research Journal*, 1-16.

Johns, T. F. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In Terence Odlin (Ed.), *Perspective on pedagogical grammar* (pp. 293—313). Cambridge: Cambridge University Press.

Johns, T. F. (1997). Contexts: The background, development and trialling a concordance-based CALL program. In A. Wichman, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100—115). London: Longman.

Johns, T. F. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp.107—117). Amsterdam and New York: Rodopi.

Kilgarriff, A. (2009). Corpora in the classroom without scaring the students. In Leung, Y. N. (Ed.), *The Proceedings of the 18th international symposium on English Teaching,* 35—45. Taipei: Crane Publishing Company Limited.

Kennedy, G. D. (1990). Collocations: The missing link in vocabulary acquisition Amongst EFL learners. *Papers and Studies in Contrastive Linguistics, 26*, 123-136.

Kettermann, B. (1995). Concordancing in English language teaching. *TELL and CALL, 4*, 4-15.

Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of prepositions: the case of Iranian EFL adult learners. *Asian EFL Journal*, *8*(4), 192-209.

Krashen, S. (1981). *The input hypothesis*. London: Longman.

Laufer, B. (1990). Words you know: How they affect the words you learn. In J. Fisiak (Ed.), *Further insights into contrastive linguistics* (pp. 573—593). Amsterdam: John

Benjamins.

Leech, G. (1994). Students' grammar, teachers' grammar, learners' grammar. In M. Bygate, A. Tonkyn, & E. Williams (Eds.) *Grammar and the Language Teacher* (pp. 17—30). London: Prentice Hall.

Lehmann, H. M., Schneider, P., & Hoffmann, S. (2000). BNCweb. In J, Kirk (Ed.), *Corpora Galore: Analysis and Techniques in Describing English* (pp. 259—266). Amsterdam: Rodopi.

Lewis, M. (1993). *The lexical approach*: The state of ELT and away forward. Hove, England: Language Teaching Publications.

Lewis, M. (2000). Teaching collocation: further developments in the lexical approach. Hove, England: Language Teaching Publications.

Li, C. C. (2005). *A study of collocational error types in ESL/EFL college learners' writing*. Unpublished master's thesis, Ming Chuan University, Taipei, Taiwan.

Lin, C. C., Chan, H. J., & Hsiao, H. S. (2011). EFL students' perceptions of learning vocabulary in a computer-supported collaborative environment. *The Turkish Online Journal of Educational Technology, 10*(2), 91-99.

Lin, Y. P. (2002). *The effects of collocation instruction on English vocabulary development of senior high students in Taiwan.* Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.

Liu, C. P. (1999 a). A study of Chinese Culture University freshman's collocational; competence: "Knowledge" as an example. *Hwa Kang Journal of English Language & Literature, 5*, 81-99.

Liu, C. P. (1999 b). An analysis of collocational errors in EFL writings. *The proceedings of the Eighth International Symposium on English Teaching*, 483-494. Taipei: Crane.

Liu, D. (2010). Is it a *chief, main, major, primary,* or *principal* concern? A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics, 15*(1), 56-87.

Liu, L. E. (2002). *A Corpus-based lexical semantic investigation of verb-noun miscollocations in Taiwan learners' English*. Unpublished master's thesis, TamkangUniversity, Taipei, Taiwan.

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quartely, 19*(2), 207-228.

Mindt, D. (1997). Corpora and the teaching of English in Germany. In A. Wichmann, S. Fligelstone, A. M. McEnery, & G. Knowles (Eds.), *Teaching and Language*

Corpora (pp. 40—50). London: Longman.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24* (2), 223-242.

O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom*. Cambridge: Cambridge University Press.

Richards. J.C., & Bohlke. D. (2012). *Four Corners Level 4*. Cambridge: Cambridge University Press.

Room, A. (1981). *Room's dictionary of distingusihables*. Boston: Routhledge & Kegan Paul.

Rutherford, W., & Smith, M. (1988).Consciousness raising and universal grammar. In Rutherford, W., & Smith, M (Eds.), *Grammar and Second Language Teacher: A Book of Readings* (pp. 107—116). Boston: Heinle and Heinle.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmitt, N. (2012). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave MacMillan.

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: Benjamins.

Sealey, A., & Thompson, P. (2007). Corpus, concordance, classification: young learners in the L1 classroom. *Language awareness, 16*(3), 208-216.

Serkan, C. (2011). Developing collocational competence through web-based concordance activities. *Novitas-ROYAL (Research on Youth and Language), 5*(2), 273-286.

Shin, S., & Wang, H. (2006). The relationship between EFL learners' depth of vocabulary knowledge and oral collocation errors. *In The 23rd International Conference on English Teaching and Learning in the Republic of China*, 964-977. Taipei, Taiwan: Jaun Tang International Publishing Ltd.

Sinclair, J. (1991). *Corpus concordance collocation*. Oxford: Oxford University Press.

Skehan, P. (2000). A Cognitive Approach to Language Learning. Oxford: Oxford University Press.

Stern, H. H. (1986). *Fundamental Concepts of Language Teaching*. Oxford: Oxford University Press.

Sun, H. C., & Wang, C. Y. (2003). Concordances in the EFL Classroom: Cognitive Approaches and Collocation Difficulty. *Computer Assisted Language Learning,*

16(1), 83-94.

Supatranont, P. (2005). *A Comparison of the Effects of the Concordance-Based and the Conventional Teaching Methods on Engineering Students' English Vocabulary Learning*. Unpublished Doctoral Dissertation. Graduate School Chulalongkorn University, Thailand.

Swan, M. (1985). A critical look at the communicative approach. *ELT Journal, 39*(1), 2-12.

Takač, V. P. (2008). *Vocabulary Learning Strategies and Foreign Language Acquisition.* British: Multilingual Matters Ltd.

Taylor, J. R. (2003). Near synonym as co-extensive categories: high and tall revised. *Language sciences, 25*(3), 263-284.

Taylor, J. R. (2007). Semantic categories of cutting and breaking: Some final thoughts. *Cognitive Linguistics, 18*(2), 331–337.

Tian, S. (2005). The impact of learning tasks and learner proficiency on the effectiveness of data-driven learning. *Journal of Pan-Pacific Association of Applied Linguistics, 9*(2), 263-75.

Tomlinson, B. (1998). Introduction. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 1—24). Cambridge: Cambridge University Press.

Touraj, T., & Zahra, F. (2012). Data-driven learning: A student-centered technique for language learning. *Theory and Practice in Language Studies, 2*(7), 1526-1531.

Tribble, C., & Jones, G. (1990). *Concordances in the classroom: A resource guide for teachers*. Essex: Longman.

Tribble, C. (1996). Corpora, concordances and ELT. *IATEFL Newsletter, 130*, 25-26.

Tseng, F. P. (2002). *A study of effects of collocation instruction on the collocational competence of senior high school students in Taiwan*. Master Thesis, National Taiwan Normal University, Taipei, Taiwan.

Tsui, A. B. M. (2005). ESL teachers' questions and corpus evidence. *International Journal of Corpus Linguistics, 10*(3), 335-356.

Vygotski, L. S. (1963). Learning and mental development at school age. In B. Simon & J. Simon (Eds.), *Educational psychology* (pp. 21—34). London: Routledge & Kegan Paul.

Wang, X. W. (2012). *The study of using data-driven learning in English grammar teaching in elementary school*. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.

Warschaucer, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly, 34*, 511-535.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*(1), 33-52.

Willis, D. (1990). *The Lexical Syllabus: A New Approach to Language Teaching*. London: HarperCollins.

Willis, D., & Willis, J. (1996). Consciousness-raising activities in the language classroom. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language Teaching* (pp. 63—76). Oxford: Heinemann.

Willis, J. (1998). Concordances in the classroom without a computer: assembling and exploiting concordances of common words. In Brian Tomlinson (Ed.), *Material development in language teaching* (pp. 44—66). Cambridge: Cambridge University Press.

Woodward, T. (1996). Paradigm shift and the language teaching profession. In J. Willis and D. Willis (Eds.), *Challenge and Change in Language Teaching* (pp. 4—9). Oxford: Heinemann.

Wu, W. S. (1996). Lexical collocations: One way to make passive vocabulary active. *Paper from the eleventh conference on English Teaching and Learning in the Republic of China* (pp. 461—480). Taipei: Crane.

Wu, W. S. (2010). The integration of corpus-based data into grammar instruction: Using *advise, recommend*, and *suggest* as an example. Retrieved from May, 10, from http://web.chu. edu.tw/ ~wswu/ publications/papers/journals/07.pdf.

Xiao, R. Z., & McEnery, A. M. (2006). *Collocation, semantic prosody and near synonymy: A cross-linguistic perspective. Applied Linguistics, 27*(1), 103-129.

Yeh, H. Y. (2013). *The effect of video production project on freshman learners' English learning motivation*. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.

Yeh,Y. Y., Liou, H. C., & Li, Y. H. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning, 20*(2), 131-152.

Yoon, H. & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2. *Journal of Second Language Writing, 13*(4), 257-283.

Young, R., & Perkins, K. (1995). Cognition and conation in second language

acquisition theory. *IRAL 33*(2), 142-164.

Zhang, X. (1993). *English collocations and their effect on the writing of native and non-native college freshman*. Publishes doctoral dissertation, Indiana University of Pennsylvania, Pennsylvan.

APPENDICES

Appendix A: Pilot study-lesson plan for corpus-based data-driven learning

Lesson Plan for Data-Driven Learning (DDL) on English near-synonyms

<table>
<tr><td>Lesson</td><td>Our-people, person, human</td><td>Class</td><td>English major freshmen in NCCU
(reading and writing I)</td><td>Class Size</td><td>13 students</td></tr>
<tr><td>Source</td><td colspan="5">Teacher-made paper-printed materials</td></tr>
<tr><td>Date</td><td>December, 20, 2012</td><td>Student</td><td>Sheng-Chi Cheng</td><td>Teacher</td><td>Siaw-Fong Chung, Ph.D.</td></tr>
<tr><td>Equipment</td><td colspan="5">paper-printed materials, notebook, projector, dictionaries</td></tr>
<tr><td>Teaching aid</td><td colspan="5">PPT slides, concordance printed teaching materials, handout, quiz</td></tr>
<tr><td>Eight Teaching Objectives</td><td>vocabulary</td><td colspan="4">1. Students will know the uses and usages among three vocabulary.
2. Students will learn the connotation and denotation meanings of three words- people, person and human.
3. Students will know some important American idioms and phrases on each word.
4. Students will know the most frequency collocations of each word.</td></tr>
<tr><td></td><td>Reading</td><td colspan="4">1. Students will read the concordance (the given context) and then try to decode the meanings of each phrase and idiom.</td></tr>
<tr><td></td><td>Writing</td><td colspan="4">1. Students can use these three vocabulary to make correct sentences.</td></tr>
<tr><td></td><td>Grammar Sentence Patterns</td><td colspan="4">1. Students can produce correct collocations between vocabulary.
2. Students know the concept of plural forms of those three words.</td></tr>
</table>

TEACHING PROCEDURE

Class Period: 45 mins

Procedures	Times	Activities	Remarks
1. Warm-up	5	1. pre-test 2. Teacher asks students three questions before listening the song [Where Is The Love? / The Black Eyed Peas http://www.youtube.com/watch?v=WpYeekQkAdc (4:13)] (DET: Guess from textual context) (MET :Use English-language media) (a) What is the difference between *people*, *person* and *human*? (b) In your opinion, which context is more suitable for *people/ person/human*? (SOC:Ask classmates for meaning) (MEM:image word's meaning) 3. Teacher gives lyric handouts[**handout 1**] to students. Students are required to pay attention to the usage of *people* and *human* when they are listening to the song. 4. After listening the music, students discuss their findings and share with each other. 5. Teacher generates students' answers.	notebook PPT lyric handout
2. Presentation	20	Students learn the three words and their usages, collocations, phrases and idioms through paper-printed concordances which prepared by the teacher before the class. 1. Teacher distributes six-page paper-printed idiom and phrases concordances **[handout 2]** to student and also a worksheet to students. 2. Teacher divides students into three groups (two groups are four people and one is five people) (SOC: Study and practice meaning in a group)	notebook PPT paper-printed concordance worksheet

		3. In the beginning, the teacher introduces the concept of *near-synonym*, *corpus* and *concordance* and then demonstrates the procedures of observing the concordance line. 4. The teacher takes one of the idioms *boat people* as an example to emphasizes students on their observing focuses. 5. Teacher asks students to write down their observations on the worksheet and then discusses answers in order to make sure learners recognize the concept of idioms and phrases. (SOC: Discover new meaning through group activity)	
3. Practice	15	Corpus-based research on three vocabulary by investigating concordance lines individually: *people*, *person*, *human*. The steps in the practice section are followed by students: (1) Students observe the concordance line and then write down their findings in three aspects within groups. (COG: Keep a vocabulary notebook/ Take notes in class) (MEM: Learn the words of an idiom together) (MEM: connect the word to its synonyms) (DET: Guess from textual context) (SOC: Discover new meaning through group activity) (2) Students discuss their findings and then try to guess the word meaning and write down in the worksheet. (MEM: image word's meaning) (3) Teacher will ask each group to talk about their findings and then generalize the results. If the teacher	notebook PPT worksheet paper-printed concordance projector

		notice that there are some missing parts which students make mistakes, the teacher will supply the correct answers. Students are encouraged to ask teacher questions. (SOC: Ask teacher for paraphrase or synonym of new word) (4) In each word, the observation sequences are shown in following: (a) Try to identify the meaning of each *idiom* through the paper-printed corpus teaching materials and then infer the meaning within the context. (b) Teacher will show answers and then talk about correct usages of each task. Task 1: people 1. Phrase: *boat people* Task 2: person 1. phrase: *lay person, people person, in the person of, be one's own person* 2. idiom: *snake oil salesperson* Task 3: human 1. phrase: *complete human, human person*	
4. Wrap up	5	1. Teacher reviews the core meanings of three vocabulary separately by giving students a handout of comparisons of five on-line dictionaries.**[handout 3]** (DET: Monolingual dictionary) 2. Review the *idiom* and *phrases* of each target word. 3. The students have a post-test.**[handout 4]** (MET: Testing oneself with word tests) 4. The students have a questionnaire in order to explore learners' perceptions of learning in a corporal DDL approach.**[handout 5]**	Notebook PPT worksheet Quiz questionnaire

Appendix B: concordance lines handout: idioms and phrases (experimental group)

1. Boat people

1	Francis Maude, Minister of State in the UK Foreign Office, visited Hanoi on Feb. 18-21 in an unsuccessful attempt to persuade the Vietnamese government to accept a policy of involuntary repatriation from Hong Kong of those Vietnamese "**boat people**" screened out as "economic migrants" .
2	Following a meeting of United States, United Kingdom and Hong Kong government officials in Washington on June 6, the US government stated that it was to drop its opposition to negotiations between the UK and Vietnamese governments for the "involuntary repatriation" of Vietnamese **boat people** not considered to be genuine refugees, and to the possible establishment of an internationally managed centre in Vietnam to hold them on their return.
3	Now, after a lull of several years, the **boat people** are heading for Australia again, this time from Cambodia
4	Vietnamese-UK talks on **boat people** issue
5	Since 1945 this figure is put at 12 million and we have recently seen dramatic examples of the same desperate remedy by the **boat people** of Vietnam and those escaping from Cuba.
6	It is likely that a group of fewer than 100 **boat people** will be sent back just before the arrival in Hong Kong of the Foreign Secretary, Mr. Douglas Hurd, planned for mid-January.
7	The number repatriated in this way has amounted to tens of thousands annually since 1980, when the present rules were instituted — a figure not too far from that of the **boat people** whom Britain and Hong Kong governments expect to send home to Vietnam.

Your finding and definition:

2. Idiom: lay person

1	Before we go any further, it is important to point out here that it would be far too ambitious (and potentially risky) for the **lay person** and, indeed, the average aroma therapist, to attempt treating serious disorders with essential oils
2	This work is relevant to the topic of naive physics discussed earlier, for the physical knowledge exploited by the program is comparable to that of the **lay person** rather than the physicist
3	Although all life and pension policy providers are now required by law to state their charges in writing, it is often extremely difficult for a **lay person** to work out exactly what is involved; or to make meaningful comparisons between one organization's charges and another's.
4	It quashed the sentences imposed on those convicted in the Good Parliament, and, at the suggestion of the commons, agreed to a poll tax of 4d on every **lay person** over fourteen years of age.
5	In the **lay person**'s view, and in the view of some professional linguists, the norms of language are associated with notions of standardization and 'correctness' or with hierarchical dimensions of social structure (or all of these), and they are usually felt to be institutional : that is, they are thought of as being prescribed by authority through the writing system, the educational system and other agencies (for a relevant discussion, see J. Milroy and L. Milroy, 1985a).
6	What about the **lay person**, the man or woman in the street or in the home, will they actually use computers in a graphical sense?

3. Idiom: people person

1	Maria Panayiotis is a self-confessed '**people person**'
2	Alison is a '**people person**' and has served on the consultative committee.

4. Phrase: be one's own person

1	She, who had been forced to **be her own person**: first, by circumstances; second, by temperament; third (as she fiercely told herself) by inclination … and yet she, Jane, could not get him, Christopher, out of her mind
2	This is why so many women who have battled to '**be their own person**' live in fear of falling in love and losing their individuality.
3	he's got to **be his own person**

5. Phrase: in the person of

1	After all, Pound had married England — not figuratively, but literally, **in the person of** Dorothy Shakespeare; and Ben Hecht in 1918 reported that Pound was ‘a doting monogamist’
2	The mother church at Jerusalem in the apostolic age had a single head **in the person of** James, the Lord's brother.
3	You can also be more experimental in a story, write **in the person of** a ghost or a cat, do things you couldn't get away with in a novel.
4	Then there was a real showman **in the person of** debonair Harry Pryce who conducted ‘Stag Party’, ‘Musical Mirror’ and ‘From Leicester Square to Broadway’ for many years.
5	The English came very close to establishing a real live Queen of Beauty **in the person of** Elizabeth I; her cult was that of the ever-lovely and ever-youthful Virgin Queen.
6	Their hurt found expression **in the person of** the secretary's wife, sitting opposite.

6. Idiom: living person

1	I remember removing the name-tapes of the former wearers (who had left the school) with a savage feeling that these people had to be got rid of, otherwise they would possess me as a **living person** is possessed by a ghost.
2	But as the child could not in the very nature of things acquire rights correlative to a duty until it became by birth a living person, and as it was not until then that it could sustain injuries as a **living person**, it was, we think, at that stage that the duty arising out of the relationship was attached to the defendant, and it was at that stage that the defendant was, on the assumption that his act or omission in the driving of the car constituted a failure to take reasonable care, in breach of the duty to take reasonable care to avoid injury to the child.
3	To this question the defendant answers ‘No,’ because at the time of his neglect the plaintiff was not in existence as a **living person**, had no separate existence apart from her mother, was not capable of suing to assert a legal right, and was not a legal person to whom he could be under a duty.
4	For one thing, the Hebrews did not divide man up into spirit, mind and body as we tend to do: they thought of him as a single entity, an animated body, a **living person**.
5	Mrs. Hughes became Britain's oldest **living person** ever on 25 February, 1992.
6	There had been conflicting opinions by individual judges on whether injuries must be sustained by a **living person** before next-of-kin could sue.

Your finding and definition:

7. Idiom: complete human

1	Only the male possesses full or **complete human** nature.
2	And I wanted to live and be a whole, **complete human** being.
3	(Even nursing of the terminally sick relates to the significance of death as the end of a **complete human** life, not just to the final days in hospital.)
4	The **complete human** being has to be able to combine both these kinds of mental function, the rational-semiotic-orderly on the one hand and the sensory-emotional-disorderly on the other.
5	He describes (without offering evidence) the bust as a 'compressed substitute for the **complete human** form'.
6	But with Auerbach the premise of the bust as substitution for the **complete human** form seems untenable.
7	It was the same sensitive touch he brought to their lovemaking, and the knowledge only added to her store of love for him, for she saw how he was the sort of man who was all of a piece, mentally and physically, a **complete human** being, someone she could only admire, despite the hell he put her through.
8	We are aware that serial killers and the like are merely expressing themselves, working out their various hang-ups and generally freeing themselves from those inhibitions which might, if suppressed, make them less **complete human** beings.

Your finding and definition:

diligent (adj)

1. someone who is diligent works hard and is careful and thorough:

ex: a diligent student

2. **:** characterized by steady, earnest, and energetic effort **:**

industrious (adj)

1. someone who is industrious works hard [= hard-working]

disrupt (verb)

1. to prevent something from continuing in its usual way by causing problems:
2. to throw into disorder

ex. *agitators trying to disrupt the meeting*

ex: *Traffic was disrupted by a hoax bomb.*

ex: Climate change could disrupt the agricultural economy.

interrupt (verb)

1 [intransitive and transitive] to stop someone from continuing what they are saying or doing by suddenly speaking to them, making a noise etc:

ex. Will you stop interrupting me!

ex. Sorry to interrupt, but I need to ask you to come downstairs.

2 [transitive] to make a process or activity stop temporarily:

ex. My studies were interrupted by the war.

decision (noun)

1. [countable] a choice or judgment that you make after a period of discussion

make a decision/ take a decision

(=make an important decision, especially after considering carefully)

reach/come to a decision

a big/major decision

a difficult/hard/tough etc decision

ex. *Do you ever wonder if you **made** the right **decision**?*

*ex. No final **decision** has been **taken**, but it seems likely that the two companies could merge in the near future.*

*ex. We finally **came to a** firm **decision** on the matter.*

2. [uncountable] the quality someone has that makes them able to make choices or judgments quickly and confidently [≠ indecision]:

ex. the ability to act with speed and decision

3. [uncountable] the act of deciding something:

ex.The Court has the ultimate power of decision.

determination(noun)

1. [uncountable] the quality of trying to do something even when it is difficult

determination to do something

ex. Yuri shows great determination to learn English.

*ex. his **dogged determination*** (=very strong determination) *to succeed*

2. [uncountable and countable] *formal* the act of deciding something officially

determination of

ex. the determination of government policy

3. [countable] *formal* the act of finding the exact level, amount, or cause of something

determination of

ex. accurate determination of the temperature

resolution(noun)

1. DECISION

[countable] a formal decision or statement agreed on by a group of people, especially after a vote

ex. The resolution was passed by a two-thirds majority.

*ex. a **resolution calling for** a ban on dumping nuclear waste*

ex. They have failed to comply with the resolution.

2. SOLUTION

[singular,uncountable] when someone solves a problem, argument, or difficult situation

resolution of

ex. a forum for the resolution of commercial disputes

3. PROMISE

[countable] a promise to yourself to do something

resolution to do something

*ex. Carol **made** a **resolution** to work harder at school.*

New Year's resolution

(=a resolution made on January 1st)

4. DETERMINATION

[uncountable] strong belief and determination:

ex. Then, with sudden resolution, she stood up.

5. Clear picture [uncountable and countable]

the power of a television, camera, MICROSCOPE etc to give a clear picture

high/ low resolution

(=how clear or unclear the picture is)

6.

a **:** the process or capability of making distinguishable the individual parts of an object, closely adjacent optical images, or

sources of light

b **:** a measure of the sharpness of an image or of the fineness with which a device (as a video display, printer, or scanner) can produce or record such an image usually expressed as the total number or density of pixels in the image
ex. a *resolution* of 1200 dots per inch

Appendix D: Vocabulary Acquisition Questionnaire

各位同學:
這份問卷主要是想要了解你們對於學習第二外語單字的學習經驗。所有的資料都將會被保密，並只拿來做實驗分析。請根據你的自身經驗回答問題，非常感謝你們的合作。

國立政治大學英文系英語教學所
指導教授: 張郇慧 教授
研究生: 陳聖其

1. 姓名: _____
2. 性別: □ 男 □ 女
3. 系別/組別: _____
4. 年級: _____
5. 年紀: _____
6. 英文科指考或學測分數: 指考 ____ 級分/ 學測 ____ 級分
7. 英文學習經驗:
   □ 少於10年 □10~15 年 □15~20 年 □ 20年以上
8. 你曾經待過英語系國家嗎?
   □ 從未 □ 少於 6 個月 □ 6 個月~ 1 年 □ 1~ 3 年
   □ 其他:_____ years
9. 你的母語(第一語言) : _____
10. 整體來說，你喜歡使用電腦嗎? □ 喜歡□ 不喜歡
11. 針對個人的需求(例如:E-mail)，你使用使用電腦的頻率(選一個選項)
   □ 一天好幾次 □ 一天一次 □ 一週五次 □一週一次
   □ 一個月一次 □ 很少使用 □其他: _____
12. 針對於學校的作業，你使用電腦的頻率(選一個選項)
   □ 一天好幾次 □ 一天一次 □ 一週五次 □一週一次
   □ 一個月一次 □ 很少使用 □其他: _____
13. 當你針對個人的需求(例如: E-mail)使用電腦時，你會使用的語言是?
   □ 英文 □ 母語(中文) □ 兩者都會使用 □ 其他:_____
14. 你曾經聽過語料庫(Corpora)嗎? □ 有 □ 沒有

15. 你曾經使用過語料庫嗎（Corpora)嗎?□ 有 □ 沒有（若有請回答第 16 題）

16. 如果有使用過，哪一個語料庫是你曾經使用過的? _____

17. 在學習英文的過程中，你會使用字典來幫助自己學習嗎? □ 是 □否

18. 如果有，哪一種字典是你最常使用的?（請看過每一個選項）

□ 雙語字典（e.g.英文-中文）□ 英英字典 □ 紙本字典 □線上字典

□ 電子辭典

19. 考過的檢定

（1）全民英檢 □ 初級 □中級初試 □中級複試 □中高級初試

□中高級複試 □ 高級初試 □高級複試

（2）多益: _____ 分

（3）托福: _____分

（4）雅思: _____分

（5）其他: _____

Appendix E: Pre and post-test

English Vocabulary Knowledge Test

Name : _____

I. Multiple choices: (40%)

1. ( ) Baby tigers learn how to catch other animals if they _____ their mothers. (1) see (2) look at (3) watch (4) look for
2. ( ) Hey, what's wrong? You really _____ unhappy. (1) see (2) look (3) watch (4) look for
3. ( ) The president wanted to take his message directly to the _____ (1) people (2) persons (3) human
4. ( ) The house _____ the Nixon family more than 5 million NT dollars five years ago. (1) cost (2) paid (3) spent (4) took
5. ( ) Last Friday, our boss said he was going on a business _____ to meet some new clients. (1) tour (2) trip (3) travel
6. ( ) The differences are not statistically _____ .Now this is a sort of cost benefit analysis based on what would have happened if we'd followed these protocols ourselves. (1) important (2) eventful (3) significant (4) eventual
7. ( ) I get the personal _____ from this business (1) benefit (2) advantage.
8. ( ) The dress is a bit loose around the waist, but it shouldn't cost much to have it _____ .(1) altered (2) changed.
9. ( ) My house is_____ the harbor, in the vicinity of the Kowloon Park. (1) cross (2) across
10. ( ) He walked _____ the stage and bowed to the audience. (1) cross (2) across
11. ( ) But a new element in Montaigne's essay was its suggestion that the earthly _____ might be a still existing real place. (1) heaven (2) paradise
12. ( ) He made a/an _____ attempt to learn Russian. (1) diligent (2) industrious
13. ( ) Climate change and rising sea levels could seriously _____the agricultural economy of Southeast Asia. (1) interrupt (2) disrupt
14. ( ) The National Council for Civil _____ is a non-partisan, voluntary

organization which has campaigned effectively to defend this country's _____ for over 50 years.

(1) Freedom ; liberties (2) Freedom ; freedom (3) Liberties ; freedom (4) Liberties ; liberties

15. ( ) You'd have more _____of catching the train if you took a bus to the station instead of walking. (1) chance (2) opportunity

16. ( ) I congratulate the Mr. Conway on moving a motion about employment policies. Although I am not <u>entirely</u> _____ by the contents of the motion, I agree with parts of it. (1) convinced (2) persuaded

17. ( ) There are psychological effects of fear and depression, which may significantly impair the victim's enjoyment of life long after the physical <u>wounds</u> have_____. (1) cured (2) healed.

18. ( ) Are the apples _____enough to eat yet?

(1) mature (2) big (3) ripe (4) rape

19. ( ) The Directors may <u>authorize</u> some _____ to execute and deliver on behalf of each Defaulting Member the necessary transfer(s) and the Company may receive the purchase money in trust for each of the Defaulting Members. (1) people (2) persons (3) human

II. Match each sentence with the correct word or phrase (52 %)

(動詞如果有必要，需做變化) （可以重複選）

A. join	B. attend	C. participate	D. enter

20. He did not _____ the party yesterday.

21. Will you _____ us in a game of cards?

22. All the students actively _____ the <u>thorough cleaning</u>（大掃除）.

A. continue	B. begin	C. start

23. Maybe there is something wrong with my car. I cannot _____ it.

24. As the Queen entered, the music _____ , we curtsied, <u>smiled</u> and _____ to move as we had never moved before. Out of the corner of my eye I saw the Queen watching us and talking to the Director.

25. Administration reported that ozone readings over the South Pole in mid September

were 15 per cent lower than those taken a year earlier. If this trend _____, the ozone hole this year will be the deepest ever.

A. regard	B. consider	C. think

26. The responses we received to the consultation document were _____ carefully and my officials discussed the issues raised with the auditing practices board and with other interested parties in the regulated sectors.
27. I do not _____ the prospects of the company favorably.
28. I'm very surprised about that, I really _____ , I have a big improvement on my test.

A. realize	B. comprehend	C. understand

30. One common theme is to better _____ the preconditions（先決條件）and effects of the diffusion of electronics（電子學的普及）based production technology on the structure and competitiveness（競爭力）of UK industry and or energy supply and demand.
31. I've just suddenly _____ actually that I'm the only health worker. Could I ask a question please?
32. In order to fully _____ the nature of the relationship between these sectors of the economy, it is necessary to look at the empirical material.

A. decision	B. determination	C. resolution

33. A: Look, Jeff can walk all by himself. He injured so seriously then. B: Yes, he fought the illness with courage and _____ .
34. Their final_____ is to go to the cinema.
35. My _____ is to be more open-minded and to work harder in school.

A. huge	B. big	C. large	D. enormous

36. Those who found difficulty in settling at Bunce Court had to sort out their own problems. There was no time for laggards(懶散的人). Pupils and staff were under

_____ pressure.

37. Fashion photographs in magazines make a/an _____ impact on young people's self-image (自我形象), particularly females aged 13-19.
38. Could I try these shoes in a _____ size?
39. The company had been very successful for New Zealand, which since 1983 had been able to increase by 170 per cent its exports（輸出物資）of mainly agricultural products（農產品）to the relatively _____ Australian market.

A. duplicate	B. replicate	C. copy

40. I picked up a _____ of their summer travel brochure (印成小冊子的宣傳品).
41. These findings are _____ by the recent local crime surveys, which have constituted the empirical core of the 'New Left realism' in criminology（犯罪學）.
42. New York to provide a full range of fulfillment services to software developers: translated into English, that means _____ floppy disks, producing and printing manuals, doing foreign language translation, and packaging and assembly.

A. people	B. person	C. human

43. Wolves will not usually attack _____ .
44. The clear implication is a future world _____ by fewer brokers.
45. In her imagination the flat was suddenly _____ with ghosts.
46. Guns hidden on his _____ .

III .Conversation (8%)

Complete the conversation with the correct word from the box

A. option	B. date	C. ensure	D. appointment
E. inhibit	F. assure	G. choice	H. prohibit

(in the office)

Tom: Hi, May.

May: Hi, Tom. Do you make a good 47._____ for selecting the new supplier（廠

商).

Tom: Well….. There are more than twenty suppliers bid（競標）this project. It is hard to make a decision. But, I think I will choose the most famous one.

May: Do you mean the IBM?

Tom: Yes, I have to make a/an 48. _____ with their president to 49._____ the contract is well-organized.

May: It sounds good. I have to notice that we have to finish this project before April.

Tom: Why? We don't have a lot of time to deal with this project.

May: According to our company procedure, when we find the new supplier, we have to run the personnel procedure before meeting the supplier.

Tom: OK! I know. The regulation（規定）of our company 50. _____ the procedure of running a new project.

Appendix F

語料庫為本之近似詞教學學習經驗問卷（後測問卷）

各位同學你好:

在參加完八週的語料觀察學習法後,為了瞭解大家的學習態度以及對語料觀察學習整體教學的看法，請大家根據自己的學習活動經驗填寫這份問卷，問卷內容包含了對語料觀察學習法的學習態度、對語料觀察導向教學活動的回應以及對老師的教學感知。老師希望能聽取您寶貴的意見，所有選項沒有對錯之分，您的意見會受到尊重以及保密，所收錄的數據只供學術研究之用，與此科目成績無任何關係，請依照自己的想法和意見放心作答!

感謝你的合作!

第一部分:語料觀察學習態度

圈選題(以下問題請依照程度作答)

	非常同意	同意	不同意	非常不同意
1. 我喜歡透過電腦來學習英文。	4	3	2	1
2. 透過電腦的輔助進行英語學習活動，對我學習英文有很大的幫助。	4	3	2	1
3. 透過語料觀察學習配合老師整理過的詞彙彙編索引(紙本語料庫)來學習英文單字，可以提升我對英語學習的興趣。	4	3	2	1
4. 透過語料庫的資料來學習單字，在未來可以讓我避免許多用字上的錯誤。	4	3	2	1
5. 透過語料觀察學習來學習單字，讓我上課更加專心。	4	3	2	1
6. 這樣的教學方式，可以讓我在測驗時有好的表現。	4	3	2	1
7. 透過語料觀察學習來學習單字，讓我很有成就感。	4	3	2	1
8. 透過語料觀察學習來學習單字,我認為我的語感以及分析能力提升了。	4	3	2	1
6. 透過語料觀察學習的教學方式來學習單字，可以加深我對英文單字的記憶。	4	3	2	1
7. 透過語料觀察學習來學習單字很適合我。	4	3	2	1
8. 透過語料觀察學習配合老師整理過的詞彙彙編所引(紙本語料庫)來學習英文單字，是浪費時間。	4	3	2	1
9. 透過語料觀察學習來學習單字，讓我很焦慮。	4	3	2	1
10.整體來說，我對於運用以語料庫為基礎的英語字彙教學情境感到滿	4	3	2	1

意。(我喜歡這樣的上課方式)				

第二部分:資料驅動導向教學活動感知

圈選題(以下問題請依照程度作答)

	非常同意	同意	不同意	非常不同意
11. 透過語料觀察學習(紙本語料庫)進行的教學活動比起傳統教科書上的活動有趣多了。	4	3	2	1
12. 透過語料觀察學習(紙本語料庫)進行的教學活動會分散我對英文字的注意力。	4	3	2	1
13. 如果有機會，我希望後能夠再利用語料觀察學習法進行其他的教學活動，例如文法學習或是寫作訓練。	4	3	2	1
14. 我看得懂學習單(紙本語料庫)上的單字。	4	3	2	1
15. 我看得懂學習單中(紙本語料庫)的句型。	4	3	2	1
16. 我覺得學習單中(紙本語料庫)所提供的例子不夠多。	4	3	2	1
17. 我覺得和同學一起發現單字規則的過程很有趣。	4	3	2	1
18. 我覺得和同學一起發現單字規則的過程讓我很有成就感。	4	3	2	1
19. 在資料驅動學習(紙本語料庫)進行的教學活動中，我從英文較好同學身上學到如何觀察語料以及單字句子的用法。	4	3	2	1
20.(a) 我想要透過實際操作線上語料庫的方式來學習單字，而不是透過紙本語料庫的學習方式	4	3	2	1

(b) 請說明原因

_____

_____

21. 你最喜歡這堂課的什麼? 為什麼?(至少 30 字)

_____

_____

22. 你最不喜歡這堂課的什麼? 為什麼?(至少 30 字)

_____

_____

23. 對於這堂課，你覺得有哪些是需要改進的地方或是值得推薦的地方，寫下你寶貴的意見。(至少 20 個字)

_____

_____

第三部分:對老師教學感知

圈選題(以下問題請依照程度作答)

	非常同意	同意	不同意	非常不同意
24.透過老師的教學，讓我學到許多單字上的知識。	4	3	2	1
25.老師教學充滿熱忱。	4	3	2	1
26.老師備課認真。	4	3	2	1

＊問卷結束，感謝您的作答! ＊

Appendix G

English Version of learners' perception in corpus-based data-driven learning questionnaire

Please circle the number according to degree: 4 represents "strongly agree"; 3 represents "agree"; 2 represents "disagree"; 1 represents "strongly agree."

Part I. Learners' attitude toward data-driven learning

	Strongly Agree	Agree	Disagree	**Strongly Disagree**
1. I like to learn English with the use of computer technology in English class.	4	3	2	1
2. I think activities through computer technology can help me in learning English.	4	3	2	1
3. I think using corpus-based paper-printed concordances to learn vocabulary can enhance my interest in learning English.	4	3	2	1
4. Thanks to the corpus data, I think I will avoid certain errors of vocabulary usage in the future.	4	3	2	1
5. I become more focused in class when I learn English vocabulary through data-driven learning.	4	3	2	1
6. Through data-driven learning approach, I think I can get higher grades in the test.	4	3	2	1
7. I have a sense of achievement when I learn English vocabulary through data-driven learning approach.	4	3	2	1
8. I think using corpus-based paper-printed concordances to learn vocabulary promotes my language awareness and decoding ability。	4	3	2	1
9. I think vocabulary exercises through corpus-based paper-printed concordance enhances my memory for English vocabulary.	4	3	2	1
10. I think learning vocabulary with the use of corpus-based paper-printed concordance is suitable for me.	4	3	2	1
11. I think using corpus-based paper-printed concordances to learn vocabulary in English class waste time.	4	3	2	1
12. I feel anxious when I learn vocabulary through data-driven learning approach.	4	3	2	1
13. In general, I feel satisfied with the experience of learning English vocabulary in a corpus-based learning environment In that, I like the way when I learn English vocabulary.	4	3	2	1

Part II. Learners' attitude toward data-driven learning activity in language learning

	Strongly Agree	Agree	Disagree	**Strongly** Disagree
14. I think doing vocabulary exercises through paper-printed concordance are more interesting than doing identical exercises of traditional teaching on textbooks.	4	3	2	1
15. I think doing vocabulary exercises through corpus-based paper-printed concordances would distract me from English words.	4	3	2	1
16. After using data-driven learning to learn English vocabulary, in the future, I would like to do other similar corpus activities in class, such as grammar learning or English writing training.	4	3	2	1
17. I can understand the vocabulary on the course material (paper-printed concordance) in English class.	4	3	2	1
18. I can understand the sentence pattern on the course material (paper-printed concordance) in English class.	4	3	2	1
19. In the course material, I think it is not enough to provide sufficient English examples for learners.	4	3	2	1
20. In data-driven learning activity, I feel it is great for finding vocabulary rules patterns with teammates cooperatively.	4	3	2	1
21. In data-driven learning activity, I have a sense of achievement finding vocabulary rules and patterns with teammates cooperatively.	4	3	2	1
22. In data-driven learning activity, I can learn a lot of English vocabulary rules and patterns from other peers who have high proficiency in English.	4	3	2	1
23.（a）I would like to explore corpus on computer rather than through prepared paper-based exercises.	4	3	2	1

（b）Please explain the reason?

_____

_____

24. What do like the most in these lessons? Why

_____

_____

25. What don't you like in these lessons? Why?

_____

_____

26. For this class, do you think in which part is needed to be improved or it is worthy to be recommended. Please write down your opinion.

_____

_____

Part III. Learners' feelings toward teachers 'teaching

	Strongly Agree	Agree	Disagree	**Strongly** Disagree
27. I think I learn a lot of vocabulary knowledge from teacher's teaching.	4	3	2	1
28. I think the teacher with an enthusiastic attitude in teaching.	4	3	2	1
29. I think the teacher is carefully well-prepared teaching materials before class.	4	3	2	1

＊This is the end of questionnaire. Thank you for your cooperation.＊

Appendix H

Sample of group interview Questions

1. Have you ever been taught by this teaching methodology? If this is the first time for you to learn English through DDL, what is your feeling or thoughts of this teaching methodology?
2. During eight-week data-driven learning instruction, what impressed you the most?
3. For English learning English in the classroom, what are your opinions about learning with traditional teaching and learning with data-driven learning?
4. During the data-driven learning activities, what was the relationship and group atmosphere among your group?
5. Do you think DDL approach is very challenging? If yes, did you make any changes in your study after participating in DDL activities?
6. Did data-driven learning instruction make you want to learn English more?
7. Did you think you improve your English vocabulary knowledge is improved after instruction?