

**The Effects of Lexical Input on L2 Writing: A
Corpus-informed Approach**

HUANG, Zeping

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Philosophy
in
English
(Applied English Linguistics)

The Chinese University of Hong Kong

August 2010



Thesis /Assessment Committee

Professor Gerald Nelson (Chair)

Professor George Braine (Thesis Supervisor)

Professor Gwedolyn Gong (Committee Member)

Professor Antonia Chandrasegaran (External Examiner)

Acknowledgements

During my study in the Department of English at CUHK, I am indebted to many people who have helped me in many ways for successfully completing my Mphil study in Applied English Linguistics.

First, my heartfelt gratitude goes to my supervisor Professor George Braine. From thesis writing to oral examination, he constantly offered me insightful feedback and guidance; among these, his recommendation of Dr. John Milton's *Word Neighbors* and *TOFEL TWE* as reliable tools for my data processing has greatly consolidated my research. His patience, kindness and understanding smoothed my path in different stages of my study. I would like to say, I feel much honored to be Professor Braine's student, his last Mphil student at CUHK.

I am also grateful to Professor Gerald Nelson, who guided me to the field of corpus linguistics. His expertise on corpus linguistics raised my interest in this computational method of analyzing language in use from a different perspective. In my preparing a self-compiled corpus at the outset of this study, Professor Nelson offered me an opportunity to be his research assistant in the *ICELite* project. From this project, I learned an innovative method of online texts collecting, which helped shape my own corpus. He also kindly obtained for me the *LOCNESS* corpus, allowing me to enrich the language data of my own corpus. I am much appreciated for his careful reading of this thesis. His useful comments provided me more insights into corpus linguistics.

I would also like to express my sincere gratitude to Professor Gwendolyn Gong for her kindness, understanding and support all along my study at CUHK. Her invaluable feedback and encouraging comments have been like a beacon leading me to the right track in conducting this research. Her concerns and caring for my family, especially my little daughter, relieved the stressful side of my postgraduate life.

My appreciation also goes to my external examiner, Professor Antonia Chandrasegaran, for her insightful and encouraging comments and suggestions. Here, I am also thankful to Professor Peter Skehan, Professor Jett Hansen and Ms. Tracy Liang for their kind help with my temporary leave in 2008.

I also wish to acknowledge my former colleague and friend Phillip Callahan. As a serious writing rater, he devoted hours to assessing my data, and assisted me in finding another two IELTS examiners to be my second and third writing raters. My appreciation is also extended to the students participating in this study for their cooperation and enthusiasm in learning English writing.

My special thanks go to my friends and fellow students, Dai Binbin, for her always being at my side sharing tears and laughter of our postgraduate life; Ren Hongtao for his inspiring discussions and being an older brother encouraging me when I encountered difficulties in my research life; Bruce Jiang and Gavin Bei, for their attentiveness to queries and useful suggestions and comments.

My deepest appreciation goes to my family. I am indebted to my parents-in-law, especially my mother-in-law for her kindly taking care of my one-year-old daughter when I was away from home for my study. In addition, my hearty appreciation goes to my mum and dad for their always supporting me to pursue a higher degree.

Last but not least, I would like to say thanks to my beloved husband who has been my soul mate and best friend for years. His genuine love and unfailing support reinforced my mind to pursue an intellectual life. His patience in discussing this study helped shape my research design. At the final stage of my thesis writing, he also helped me go through the whole thesis and do the tedious editing work. Finally, I would like to thank God, for endowing me an unexpected but cherish gift, my dearest daughter, to be my spiritual crutch.

內容摘要 (英文)

Abstract

Abstract of thesis entitled: The Effect of Lexical Input on L2 Writing: A Corpus-informed Approach

Submitted by.....黃澤萍 (HUANG Zeping)

For the degree of.....哲學碩士 (Master of Philosophy)

at The Chinese University of Hong Kong in July, 2010

Based on Tribble's (2002) framework of corpus-informed writing instruction, the current study adopted empirical research methods to investigate whether and to what extent concordance learning materials make an impact on L2 learners' argumentative writing.

A topic-specific corpus was established at the outset to develop KWIC (Key Word In Context) concordance learning materials. As a controlled experimental study, forty third-year Chinese EFL undergraduates were enlisted and randomly assigned to a control group and an experimental group. Lexical input as an independent variable was tested with students in both groups by requiring them to produce writing using a prescribed list of words consisting of two types, i.e., signalling nouns (words frequently used in the argumentative genre), and content schemata nouns (words related to the writing topic). At the pre-writing stage, the experimental group was given the target words supplied with corpus concordances, whereas the control group the words only (the latter could consult dictionaries if necessary).

The written texts of the pretest, immediate posttest and delayed posttest were

investigated and compared within and between groups. Although no statistically significant difference in the overall writing quality was found between the two groups, the experimental group made better mean improvement than the control group in the two post-tests. As for the use of the target nouns, the experimental group made fewer errors and had a higher variety of collocational and syntactic patterns than the control group. Analysis of post-experiment questionnaires indicated that overall, the students had a positive attitude towards corpus-formed learning materials and found them most useful for collocational learning. However, as revealed from the learning journals, quite a number of students seemed to have an unfavourable response to content schemata input due to the cut-off concordance lines and culture-loaded information.

Based on the hard evidence found in this empirical study, it is argued that the corpus-informed writing approach could be an innovative and effective means of complementing the traditional L2 writing approaches for improving L2 learners' lexico-grammatical patterns.

內容摘要 (中文)

Abstract (Chinese)

根據 Tribble (2002)所提出的語料庫輔助寫作教學的理論模式,本研究採用實驗方法詳盡探討語料庫檢索學習材料對二語學習者英語議論文寫作所產生的作用。

本實驗通過建立一個相同主題的小型語料庫,並利用該庫設計一套語料庫檢索詞彙練習材料。該材料包含兩類詞彙,一類為指示名詞 (signalling nouns), 另一類為內容圖式名詞 (content schemata nouns)。受試者為 40 名英文水準相當的中國學生,隨機分配為控制組和實驗組。寫作實驗開始前,分別給予兩組學生同一組單詞進行學習與頭腦風暴 (Brainstorming)。實驗組通過語料庫檢索材料學習該組單詞,控制組則通過查閱字典進行學習。寫作實驗中,兩組學生根據要求在命題作文中使用該組單詞。

通過對前測,後測及延後測寫作文本組內和組間的比較,研究發現,兩組間的寫作總體質量沒有顯著差別。值得一提的是,實驗組在兩次後測中,作文平均分提高幅度均比控制組大。並且,在所測試單詞的遣詞造句方面,實驗組誤用的次數明顯比控制組少,詞語搭配及句式也更加豐富多樣。實驗後問卷調查及學生的學習日誌分析結果表明,受試者對語料庫輔助學習材料評價持正面態度,並一致認為該學習模式對習得英文詞語搭配幫助最大;然而,學生對內容圖式詞的輸入是否有助於提高寫作內容則意見不一,大部分學生認為不完整的檢索句子及其句中包含的文化負載詞對整體內容的理解造成負面影響。

以上發現在某種程度上證明語料庫輔助寫作教學方法可以有效地輔助傳統的寫作教學模式,從而幫助二語學習者更好的習得英文詞語搭配及句式。

Table Of Contents

Acknowledgements	i
Abstract.....	iii
Abstract (Chinese).....	v
Table Of Contents.....	vi
List of Tables	ix
List of Figures and Graphs	x
CHAPTER ONE INTRODUCTION.....	1
1.1. Motivation.....	1
1.1.2. The importance of language use in L2 writing	1
1.1.2. The possibilities of integrating corpora into L2 writing instruction	2
1.1.3. The need for corpus-informed approach	2
1.2. Purpose of this study.....	4
1.3. Research questions.....	5
1.4. Overall research methods.....	5
1.5. Significance of the study.....	5
1.6. Organization of the thesis	6
CHAPTER TWO LITERATURE REVIEW	8
2.1. Research on corpora and L2 writing.....	8
2.1.1. Studies on corpus use from teachers' perspective	9
2.1.2. Studies on students' direct use of corpus	10
2.1.3. Empirical Studies on corpus and vocabulary learning	17
2.2. Evaluations of the studies under review	19
2.2.1. Evaluation of research on corpus-informed teaching materials development	19
2.2.2. Evaluations of empirical research on students' direct use of corpus.....	21
2.3. Call for further studies	22
CHAPTER THREE METHODOLOGY	24
3.1. Participants	24
3.2. Research Setting	25
3.3. Materials	25
3.3.1. Corpora used	26
3.3.2. Selecting the target words	28
3.3.3. Sifting the concordance lines	29
3.3.4. Formulating the queries	30

3.4. Research design	32
3.4.1. Pretest.....	33
3.4.2. Immediate posttest	33
3.4.3. Delayed posttest	34
3.5. Procedures.....	35
3.6. Instruments	37
3.6.1. Questionnaires.....	37
3.6.2. Learning journals	38
3.6.3. Uptake sheets	38
3.7. Data collection and analysis	38
3.7.1. Holistic scoring	39
3.7.2. Analysis of the use of target words	40
3.7.3. Questionnaire responses.....	42
3.8. Chapter summary	43
 CHAPTER FOUR RESULTS	 44
4.1. Holistic scores.....	44
4.2. Use of signaling nouns (SNs)	46
4.2.1. Accuracy	46
4.2.2. Complexity.....	61
4.2.3. Retention of the target patterns	73
4.3. Content Schemata nouns.....	74
4.4. Evaluation of the concordance exercises	75
4.4.1. Effects on vocabulary learning	75
4.4.2. Effect on L2 writing.....	78
4.4.3. Difficulties in doing the concordance exercises.....	80
4.5. Chapter summary	84
 CHAPTER FIVE DISCUSSION.....	 85
5.1. Did the corpus-informed approach improve students' overall writing quality?	85
5.1.1. Cut-off sentences.....	87
5.1.2. Culture-loaded information in concordance lines	88
5.2. Did the corpus-informed approach improve vocabulary use in students' writing?	90
5.2.1. Interface of lexis and syntax	91
5.2.2. Encouraging usage-based learning.....	95
5.2.3. Raising learner awareness of collocation and colligation	97
5.2.4. Retention of lexico-grammatical patterns	98
5.3. Did students think that corpus-informed approach helped their writing?.....	100
5.4. Towards a tentative model of corpus-informed writing instruction.....	102
5.4.1. Preparing Materials.....	104
5.4.2. During the exploration of a topic-specific corpus.....	105
5.4.3. Follow-up activities after exploration of the topic-specific corpus.....	106
5.5. Chapter summary	108

CHAPTER SIX CONCLUSION.....	109
6.1. Summary of this study	109
6.1.1. Enhancement of lexico-grammatical patterns.....	109
6.1.2. Enhanced awareness of the importance of collocations.....	111
6.1.3. Pivotal role of prior grammatical knowledge in corpus-informed learning.....	111
6.1.4. Insignificant correlation between learning CSNs and ideas development.....	113
6.2. Pedagogical implications	113
6.2.1. Writing materials development	114
6.2.2. Implementation of corpus-informed activities	115
6.3. Limitations and suggestions.....	117
6.3.1. A longer experimental time frame.....	117
6.3.2. More lexical input	118
6.3.3. More comparison groups	118
6.3.4. Different proficiency levels.....	119
6.3.5. Web-based concordances and more follow-up learning activities	119
6.3.6. Case studies.....	120
6.4. Closing remarks	120
Bibliography	121
Appendix One Questionnaire One.....	133
Appendix Two Questionnaire Two	136
Appendix Three Learning Journal	139
Appendix Four Pre-writing Vocabulary Study.....	140
Appendix Five Pretest Writing Task.....	153
Appendix Six Immediate Posttest Writing Task	154
Appendix Seven Delayed Posttest Writing Task.....	155

List of Tables

Table 2.1. Empirical studies on students' direct use of corpus	11
Table 3.1. Procedures of the research	36
Table 3.2. Students' written corpora	40
Table 4.1. Descriptive statistics of the writing tests	45
Table 4.2. Appropriacy scale	47
Table 4.3. Error-free ratios in the pre-test and the immediate posttest	48
Table 4.4. Errors distributions of the experimental group	54
Table 4.5. Description of sns usage change between the pretest and the immediate posttest..	55
Table 4.6. Positive change of <i>controversy</i> in the experimental group.	56
Table 4.7. Positive change of <i>objection</i> in the experimental group	57
Table 4.8. Positive changes of <i>criticism</i> in the experimental group	58
Table 4.9. No changes in the experimental group.....	60
Table 4.10. Grammatical patterns of <i>controversy</i> in the immediate posttest.....	63
Table 4.11. Grammatical patterns of <i>objection</i> in the immediate posttest.....	64
Table 4.12. Grammatical patterns of <i>criticism</i> in the immediate posttest.....	66
Table 4.13. Grammatical patterns of <i>situation</i> in the immediate posttest.....	67
Table 4.14. Patterns of <i>premodifier</i> + <i>SN</i> in the immediate posttest between the two groups	71
Table 4.15. Patterns of <i>SN</i> + <i>postmodifier</i> in the immediate posttest between the two groups	72
Table 4.16. Occurrences and errors of the target words in the delayed posttest.....	73
Table 4.17. Effects on vocabulary learning (n=20)	76
Table 4.18. Effects on L2 writing (n=20)	78
Table 4.19. Problems in doing the concordance exercises (n=20).....	80
Table 4.20. Overall evaluation on the concordance exercises	82

List of Figures and Graphs

Figure 3.1. Excerpt of the concordance exercises in the pilot study	30
Figure 3.2. Research design.....	32
Figure 3.3. Flowchart of the experiment	32
Figure 5.1. Effects of corpus-informed approach on vocabulary use in L2 writing.....	100
Figure 5.2. A tentative model of corpus-informed approach to L2 writing	107
Graph 4.1. Mean scores of the three writing tests	45
Graph 4.2: change of the use of the SNs in the experimental group	56
Graph 4.3. Distribution of grammatical patterns in the immediate posttest.....	69
Graph 4.4. Distribution of <i>premodifier</i> + <i>SN</i> in the immediate posttest between the two groups	70
Graph 4.5. Distribution of <i>SN</i> + <i>postmodifier</i> in the immediate posttest between the two groups	70
Graph 4.6. Distribution of content schemata nouns in the immediate posttest.....	74
Graph 4.7. Distribution of content schemata nouns in the delayed posttest.....	75

CHAPTER ONE INTRODUCTION

1.1. Motivation

1.1.2. *The importance of language use in L2 writing*

Since the 1980s, the planning-writing-reviewing model developed by Flower and Hayes (1981) has been the most widely adopted model in L2 writing instruction (cited from Hyland, 2003). It is crucial for teachers to guide learners through a writing process, setting up activities for encouraging brainstorming, generating ideas about content, organizing essay structures and requiring multiple drafts. However, a lack of vocabulary has been considered the major factor that makes writing in a foreign language most difficult (Leki & Carson, 1994). As Yoon (2008) noted,

While recent writing research mainly addresses the global or discourse aspects of writing, language issues cannot be ignored. For L2 writers, the mastery of lexical and grammatical accuracy can contribute to an increased confidence in themselves as writers as well as a possible increase in the quality of their writing. (p. 45)

As a writing teacher in a Chinese EFL context, I also found that vocabulary use was one of the major difficulties for learners despite the adoption of the process and genre approaches in the writing classroom. Students often complained of having few ideas about the prescribed writing topics and of lacking sufficient vocabulary to express their views in English. In order to solve these problems, teachers would distribute some articles related to the topic to help the students come up with more ideas and expressions, but many students often ended up copying the whole sentence, or even the whole paragraph directly from the given texts. Also, student essays contained some miscollocations and broken sentences

which adversely affected language fluency and complexity. In view of the above problems, it is believed that a pedagogy that facilitates learners to acquire linguistic patterns and word usage can be developed and introduced to conventional writing instruction.

1.1.2. The possibilities of integrating corpora into L2 writing instruction

In the past decade, corpora have become “increasingly important for second language writing instruction as teaching has become less a practice of imparting knowledge and more of providing opportunities for learning” (Hyland, 2003, p.167). To enhance L2 learner writing performance, a corpus approach has been regarded as a viable approach for helping learners with their lexico-grammatical patterns (Coxhead & Byrd, 2007; L. Flowerdew, 2010) and organizational patterns (Tribble, 2001). Tribble and Jones (1997) discussed two possibilities of incorporating corpora into L2 writing classroom. First, teachers could examine a corpus to determine the most common words or patterns relative to a target genre, and write teaching materials based on the observed results. Second, students can be taught how to use a concordancer to explore the corpus themselves.

1.1.3. The need for corpus-informed approach

Although the use of corpus and concordancer can offer “one of the most exciting application of new technologies to L2 writing class” (Hyland, 2003, p.167) as described above in Section 1.1.2, two existing problems of corpus use in this area cannot be ignored, i.e. “the theoretical problem and the methodological problem” (Braun, 2007, p.31). From a methodological point of view, “the most widely accessible corpora have been developed

with linguistic research goals in mind” and they “are not necessarily the corpora with the most obvious pedagogical value” (Braun, 2007, p.31). It has been argued that a smaller and genre-specific corpus can overcome some of the shortcomings of “mainstream” corpora, especially with regard to size and diversity of content. But realistically, these small corpora have not travelled well beyond the institution in which they were created (Aston, 2004). From a theoretical point of view, Widdowson (2003) criticized the use of corpora stating that while language learning is concerned with discourse and the use of language in concrete communicative situations, corpora are only a collection of texts, that is, products of language use isolated from any communicative situation.

To minimize the potential shortcomings of corpora, it is pivotal to turn language corpora, the “text museum” as metaphorized by Braun (2007, p.44), into “a lively and motivating educational experience”. In view of this, when integrating corpora into language teaching, it would be more pertinent to adopt a corpus-informed rather than a corpus-based approach on the grounds that in the former, learning takes on both advantages of the traditional teacher-student interactivity and the technological benefits brought by the pedagogical application of corpus. The significance of corpus-informed approach to language teaching is also reinforced by McCarthy (2008):

[C]orpus-based materials are materials that try to be absolutely faithful to what the computer tells you about language use, whether you like it or not, whether it is useful or not (...)

Everything we do is based on what we get from the corpus; while corpus-informed is that teachers take from the corpus what they believe will fulfill students’ needs, and the teacher’s or material writer’s task is essentially to mediate corpus information, filtering it for pedagogical purposes. (p.569)

1.2. Purpose of this study

A number of corpus linguists have discussed the use of corpora in language learning by L2 learners. Among these, some measured the effects of corpus use in receptive vocabulary acquisition (Cobb, 1997; Steven, 1991); some focused on the students' evaluations of corpus use in writing (e.g., Yoon & Hirvela, 2004); some investigated the effects of using corpus by learners as a consultation skill for error corrections in their writing (Chambers & O'Sullivan 2004; O'Sullivan & Chamber, 2006); and others discussed the significance of using a corpus linguistic approach for preparing vocabulary and grammar of academic prose for learners in order to improve their academic writing (Coxhead, 2008; Coxhead & Byrd, 2007; Thurstun & Candlin, 1997).

In direct contrast, there has been comparatively little research that provided empirical support for the effectiveness of the corpus approach for enhancing learners' lexical input, and the extent to which it can improve learners' lexical use in writing. Apart from this, no relevant studies have measured learners' outcomes of productive vocabulary in their written products. In view of this research gap, the present study sets out to investigate the effects of corpus-informed approach on learners' L2 written products. To be more specific, the subordinate objectives of this study are threefold: 1) to examine how corpus-informed materials affect lexico-grammatical patterns of target words; 2) to explore whether corpus-informed lexical input help learners develop ideas for L2 writing; 3) to investigate learners' perceptions of a corpus-informed approach to L2 writing.

1.3. Research questions

Three research questions are formulated as follows:

1. Will a corpus-informed approach help L2 learners improve their overall writing quality?
2. Will a corpus-informed approach affect L2 learners' vocabulary use in their writing?
3. Will L2 learners think that a corpus-informed approach helped their L2 writing?

1.4. Overall research methods

This study adopted a pretest-posttest experimental design, i.e., 40 students with the same proficiency level were divided into two groups (a control group and an experimental group) who took a pretest, an immediate posttest and a delayed posttest under controlled conditions in the context of writing workshops. The written products of these two groups were compared and analyzed both quantitatively and qualitatively. In order to obtain multiple data, this study also included questionnaires and learning journals to elicit learners' evaluations and perceptions of the use of corpus-informed materials in L2 writing. The triangulation of data resources provided more reliable and comprehensive accounts on the topic under investigation.

1.5. Significance of the study

As stated below, the significance of this study can be twofold.

First, it is, so far, one of the first empirical studies investigating the effectiveness of corpus use in L2 writing by focusing on textual analysis of learners' written products.

Although research on the application of corpus to language teaching has increased in recent

years, most previous studies appear to be theoretically-based and descriptive with a focus on how to incorporate corpus into L2 writing (e.g., Bernardini, 2002; Coxhead & Byrd, 2007; L. Flowerdew, 2010; Lee & Swales, 2006; Thurstun & Candlin, 1998; Tribble, 1997, 2001, 2002). This type of research has provided valuable frameworks and detailed accounts on how corpora can be utilized to enrich learning resources in L2 writing instructions or used as a reference tool for enhancing writing skills. However, most of them only explore L2 learners' writing experience and perceptions of corpus use in the process of drafting rather than focus on examining writing products in relation to corpus use. In comparison with previous studies, this study is different because it is grounded in authentic data of students' written output and the findings are obtained from a carefully designed comparative experiment. It is, therefore, anticipated that the findings of this study will substantiate the claim regarding the usefulness of corpus in L2 writing with actual empirical evidence.

Second, by investigating the overall writing quality and lexico-grammatical patterns of the target words in students' written texts, this study not only identified the ways in which corpus-informed approach helped learners with their L2 writing, but also characterized the problems and difficulties that learners may confront during the concordance observation activities. The findings of this research can shed light on the implementation of corpus-informed activities in the actual L2 writing context.

1.6. Organization of the thesis

This thesis consists of six chapters. Chapter One presents an overview of the current

research with a brief description of research motivation, research purpose, research questions, overall research methodology and the research significance.

Chapter Two is a review of literature on corpus use in L2 writing and vocabulary learning. After scrutinizing the relevant previous studies, research gaps are summarized to indicate the necessity of this research.

Chapter Three describes the components of research methodology, including the participants, the development of corpus-informed materials, research setting, research design, research instrument, procedures, data collection and data analysis.

Chapter Four reports on the holistic scores of students' written output in the pretest, the immediate posttest and the delayed posttest, the actual use of the target nouns in both groups' written texts and students' evaluations of corpus-informed materials.

Chapter Five discusses the major findings by referring back to the research questions and previous theories in second language acquisition.

The last chapter, Chapter Six concludes the current study with a summary of the significant findings, a description of its pedagogical implications, limitations and recommendation for future research.

CHAPTER TWO LITERATURE REVIEW

This chapter reviews previous research on applying language corpora to L2 teaching in an L2 writing context conducted within the recent 30 years. Section 2.1 summarizes the relevant studies by classifying them into two categories, i.e., teachers' pedagogical use of corpus (Section 2.1.1) and students' direct use of corpus (Section 2.1.2). Research on corpus use for vocabulary teaching was also reviewed since vocabulary learning and writing were two inter-related areas (See Section 2.1.3). Section 2.2 is a comprehensive and critical evaluation of the research under review. Section 2.3 states the research gap by relating to the topic under investigation and indicates the significance and necessity of this study.

2.1. Research on corpora and L2 writing

For more than two decades, corpus has been regularly described as one of the most promising ideas in language learning and teaching (Johns, 1986; Sinclair & Carter, 2004). There have been quite a number of studies focusing on corpus application to L2 writing instructions, which can be seen in L. Flowerdew (2010). In order to keep the research scope within reasonable limit, this chapter will only consider studies concerning two perspectives, i.e., corpus use as viewed from teachers' perspective and from learners' perspective.

2.1.1. Studies on corpus use from teachers' perspective

Thurstun and Candlin (1997), by utilizing a specialized corpus, designed a workbook for students' essay writing. The workbook provided both NS (English native speakers) and NNS (Non English native speakers) university students with intensive exposure to some of the most important words in academic English. It introduced various rhetorical functions of academic essays (such as stating the topic, referring to the literature, drawing conclusion) and presented sets of concordance-based exercises on a number of the most common vocabulary items used for carrying out these functions. The selection of the vocabulary was based on a specialized corpus of professional academic writing, an electronic collection of academic texts and papers from a range of disciplines, with a total word count of over one million words. Thurstun and Candlin first identified words in University Word List developed by Nation (1990) that can be used in the performance of the rhetorical functions of academic texts. Subsequently, they used the Microconcordance program to produce sets of key word in context (KWIC) concordance lines in order to observe frequencies of use as well as the lexical patterns around these words. Finally, they extracted those lines that represented the most common collocational features surrounding the word that had been searched for.

Grounded on discourse and genre analytic frameworks, Tribble (2002) outlined "ways in which appropriate corpus resources may help learners to develop competence as writers within specific academic domains" (p. 131). The author demonstrated how to use keywords and frequency list to identify lexico-grammatical features of the text. "By examining a set of words that are statistically prominent in comparison with a general population of texts,

along with a small number of high frequency words” (p.142), Tribble concluded that it was possible to “identify patterns in the text which contribute to the special identity of the text”(p. 142).

Another two representative studies on applying corpus linguistics to teaching material development are Coxhead (2000) and Coxhead and Byre (2007). With the aid of a corpus linguistics approach, Coxhead (2000) generated an academic wordlist (AWL) of 570 headwords and 3000 words altogether. The idea of AWL was to provide a shortcut for learners to expand their vocabulary by learning the most frequently-used lexis for academic writing. In 2007, Coxhead and Byre delineated a possible way, through integrating concordances with the keyword list, to prepare teachers for teaching vocabulary and grammar for academic prose. In comparison to the traditional method of selecting and analyzing samples of academic prose, the authors stated that the web-based corpus approach, concordancing in particular, was more innovative and effective for helping teachers with material development and providing them with information about academic language. They argued that using such learning materials could benefit students to obtain the skill and knowledge needed to become effective learners of new words and their associated grammar.

2.1.2. Studies on students' direct use of corpus

The focus of academic discussions and practical corpus application has shifted from the indirect uses by teachers to the direct uses by learners. The literature on learners' use of corpus in language learning has witnessed an increase in the past two decades. These

studies can be placed into three categories according to the different writing stages on which the reviewed studies focus (See Table 2.1).

Table 2.1. Empirical studies on students' direct use of corpus

Writing stage	Previous research	Aims
Prewriting	Thurstun and Candlin (1998)	Described the corpus-based exercises for helping learners with basic academic vocabulary and evaluates students' attitudes towards the corpus learning activity
	Weber (2001)	Reported on a project which aimed at helping law undergraduates to write formal legal essays by adopting a concordance- and genre-based approach.
	Lee and Swales (2006)	Presented an overall view of an experimental corpus-based EAP course for doctoral students, aiming at raising their language awareness of academic English in their own fields.
During writing	Yoon and Hirvela (2004)	Evaluated learners' attitudes towards using an online general corpus when composing essays.
	Yoon (2008)	Reported on a qualitative study investigating the changes in students' writing process associated with corpus use over an extended period of time
Post-writing	Gaskell and Cobb (2004)	Evaluated 20 lower-intermediate L2 writers' attitudes and performance on error corrections by using computer corpus and concordance.
	Chambers and O'Sullivan (2004)	Reported on the first phase of their project of improving students' writing performance with corpus consultation.
	O'Sullivan and Chambers (2006)	Reported on the second phase of the research project of using corpus to improve language use in students' writing.

2.1.2.1. Prewriting

In Thurstun and Candlin (1998), students were instructed to learn academic vocabulary by adopting an innovative approach with a concordancing program. Students were first presented with multiple examples of the vocabulary items in context. The student examined these one-line concordances, highlighted word groups surrounding the key word to discover how they were used, and then answered questions on their use in context. Students were encouraged to become sensitive to the patterns of language surrounding the key word. The exercises following the concordances lines were designed to develop awareness of the appropriate preposition and the linking of the key word with its anaphoric referent (by the creation of lists). Thurstun and Candlin reported on the students' positive reactions on the online concordance-based vocabulary exercises, which confirmed that learning the basic words for academic written texts was beneficial for students' own academic writing. The researchers also pointed out that over-exposure to concordance lines could conceivably tire students if teaching of this type depended solely on deduction from concordance lines. Hence, a variety of different activities were provided, all based on concordanced texts but demanded different problem-solving techniques from the students.

Weber (2001) described a project which aimed at teaching NNS law undergraduates to write formal legal essays. Before writing, the students were directed to identify some structural characteristics of legal essays by working on the mini-corpora individually or in small groups. They were then asked to read through the essays again, marking each occurrence of each of the identified structural parts. In the next step, the students were asked to use concordances to check any lexical items which seemed to correlate with the

generic structures. Finally, they were asked to write mini-essays incorporating the generic features and lexical items identified in the exploration of genre-based mini corpora. Weber drew a conclusion that although the students found working with concordances a challenging task, the described concordance- and genre-informed approach assisted students to write acceptable essays.

Lee and Swales (2006) designed a corpus-based course for doctoral students in a university in the United States. Four doctorate students attended this selective course for 13 weeks. The students had a weekly 2-hour laboratory session (hands-on concordancing and EAP lessons) for the first 10 weeks, using Wordsmith (corpus analysis software) to explore language patterns and observe concordances to examine errors and raise awareness of lexico-grammar, discourse conventions and semantic prosody. Optional weekly individual consultations were also available for the students for getting comments and corrections on their written work. In Weeks 11 and 12, the students compiled their own writing corpora, i.e., one of published academic papers; the other corpus was their own writing consisting of terms papers, dissertation drafts and unedited journal drafts. In the final week, the students gave presentations of the final project based on their self-compiled corpora, presenting their discoveries with some discussions of how they felt their rhetorical consciousness was raised and reflections on what further use they might be making of corpus linguistic techniques in their future careers. Lee and Swales argued that the corpus approach allowed non-native speakers a chance to make their own discoveries instead of relying on native-speaker intuitions or grammar/style books; it also involved texts from a variety of

different writers instead of just one native speaker teacher standing at the front of the classroom.

2.1.2.2. During writing

Yoon and Hirvela (2004), which has earned praise for its thoroughness, presented a sophisticated survey on learners' attitudes towards corpus use as consultation resources. In their survey, eight intermediate students and fifteen advanced students from the language center of an American university were instructed to use *COBUILD* corpus, an online general corpus, to consult linguistic problems encountered during the process of writing. Based on the analysis of the questionnaire and interview data, the authors reported that students found corpus useful for vocabulary and grammar consultation and the development of academic writing skills. Overall, the students perceived that corpus approach was beneficial to the development of L2 writing skill and increased confidence toward L2 writing.

Based on case studies with six L2 writers in a graduate-level advanced ESL academic writing course, Yoon (2008) investigated these six students' L2 writing process with the consultation of a free online corpus, i.e., the *Collins COBUILD Corpus*. Through classroom observation, interview data analysis and examination of recall protocols, corpus search logs, class corpus search assignments and written reflections on corpus use, it was found that corpus use not only had an immediate effect for helping the students solve immediate language problems, but also promoted their perceptions of lexico-grammar and language awareness. Once the corpus approach was introduced to the writing process, the students

assumed more responsibility for their writing and became more independent writers, and their confidence in writing increased. This study identified a wide variety of individual experiences and learning contexts that were involved in deciding the levels of the students' willingness and success in using corpora.

2.1.2.3. Post-writing

Gaskell and Cobb (2004) investigated 20 lower-intermediate adult L2 writers' attitudes towards and performance in error corrections by using computer corpus and concordances. In their study, the participants attended a course which followed a process approach consisting of 10 assignments over a 15-week semester. Assignments were completed in two-week cycles, with a first draft and peer feedback in Week 1, and revision and electronic submission in Week 2. The instructor then gave feedback to each student's assignment, including online concordance links for five typical grammatical errors per week, such as the use of articles, subject-verb agreement, conjunctions, noun plurals and word order. The students were required to revise the text for final submission, and for each of the concordance-linked errors to submit a form explaining specifically what correction had been made based on what concordance information. In the last week, the researchers obtained handwritten posttest writing samples on the same topic as the pre-test samples. A comparison between the pre and post-writing samples was conducted to see if there had been any change in class-wide error rates that could be related to the use of concordancing software. Students also filled in a questionnaire asking them whether they had enjoyed the technological bent of the course and whether they felt they had gained something from

concordancing. The results showed that learners were willing to use concordances to work on grammar and they were able to make corrections based on concordances. This experiment also confirmed the practicability of the corpus approach used in the previous studies for lexical concordancing (Hanson-Smith, 1993; Johns, 1986; Johns & King, 1991; Leech & Candlin, 1986).

Another two noteworthy studies were Chambers and O'Sullivan (2004) and O'Sullivan and Chambers (2006). In 2004, Chambers and O'Sullivan investigated the effects of corpus consultation on advanced learners' writing skills. Eight postgraduate students of French wrote a 600-word essay and subsequently attempted to improve it by using concordancing software to consult a small corpus containing texts on a similar subject. To do this, they were asked to write the correction or the new words or phrases above the text of the original essay (only indicating those changes which came as a result of using the corpus). They were also directed to fill in the form which was attached with details of the search word they used, the results of the concordance and what they discovered. The analysis of the results revealed that a significant number of changes made by the learners which may be classified in order of frequency: 1) grammatical errors (gender and agreement, prepositions, verb forms/mood, negation and syntax); 2) misspellings, accents and hyphens; 3) lexico- grammatical patterning (native language interference, choice of verbs and inappropriate vocabulary) and 4) capitalization. The authors also pointed out that the students commented in their evaluation forms that corpus consultation was a quick way of checking the language issues compared with dictionary use.

After replicating the study in 2004, O'Sullivan and Chambers (2006) reported on the second phase of their research project on using corpus as consultation resource. In this study, fourteen native speakers of English at undergraduate level wrote an essay of 600 words in French. The students completed this writing task in their own time and could consult traditional language resources such as dictionaries and grammars. After they submitted their texts, the researchers underlined the errors. Learners then tried to improve the underlined errors in their texts by using an online corpus. The results revealed that the students were able to make positive grammatical changes in the proofreading of their essays by using corpus as consultation resource. Based on the outcomes obtained from these two empirical studies, the researchers concluded that concordancing software could be appreciated by students at both postgraduate and undergraduate levels. Although the postgraduate students performed slightly better than the undergraduates, both levels of students perceived that they would use corpora in their future writing.

2.1.3. Empirical Studies on corpus and vocabulary learning

Since vocabulary teaching and writing are two inter-related but not discrete areas, it is worth reviewing empirical research on corpus use for vocabulary learning. Steven (1991) is the pioneer in conducting the first controlled experiment investigating the effectiveness of learners' consultation of corpus printouts at Sultan Qaboos University in the Sultanate of Oman. In this study, a group of students were directed to recall a known word to fill a gap in a text, which was either a single gapped sentence or a set of gapped concordance. The

results of learners' performance on these two sets of exercises showed that learners retrieved words from memory more successfully when doing the concordance-based exercises.

Cobb (1997) improved Stevens (1991)'s methodology by setting up a comparative study with a control group and an experimental group in order to examine whether and the extent to which online concordance exercises were more effective for helping learners' vocabulary acquisition than the traditional vocabulary exercises. In the experiment, the subjects were first-year Arabic-speaking university students taking a year of intensive English in preparation for a full load of English-medium commerce subjects in a second year (such as accounting, marketing, and management information systems). A suite of five familiar *CALL*-type activities, with a modified concordance as its main information source, were designed and tested with more than 100 learners over an academic term. In these five activities, students accessed a 10,000-word corpus consisting of 20 texts of about 500 words, each taken from the students' reading materials. The five activities moved from easy to difficult, from word-level to text-level, and from reception to production. The activities were: 1) choosing a definition; 2) finding words; 3) spelling words; 4) choosing words for new texts; and 5) writing words for new texts. As for the control group, the students used a set of traditional vocabulary learning materials. During one academic semester learning, weekly quizzes were given to both groups. The results of the quizzes confirmed that the online concordance exercises were more effective for students' vocabulary acquisition than non-concordance exercises.

Boulton's (2010) controlled experiment, which investigated 62 low proficiency level learners in France coping with the paper-based corpus exercises and a DDL (data-driven learning) approach in comparison with the traditional teaching materials. Fifteen problematic language items were selected from students' written production. Two sets of teaching materials were distributed to the students for them to study the 15 problematic language items. One set was with concordance materials of the target items while the other with traditional materials retrieved from dictionary entries. According to the outcomes of a posttest on these 15 problematic items, the author argued that the corpus-based exercises helped students learn the language better than the traditional learning materials did.

2.2. Evaluations of the studies under review

For the purpose of evaluation, the above literature has been placed into three categories: 1) research which used corpus linguistics methods to develop teaching materials; 2) empirical research on data-driven learning.

2.2.1. Evaluation of research on corpus-informed teaching materials development

Coxhead (2000) and Coxhead and Byre (2007) conducted corpus-based studies to examine academic vocabulary, which shed lights on the analysis of vocabulary, grammar and their combination of lexico-grammatical features of a language as part of the writing class and provided substantial data on rhetorical and discourse features of academic prose. However, one of the challenges of AWL developed in Coxhead (2000) is that the wordlist was released solely as a list of individual words and their families. No indication of the context and patterning in which these words occurred was presented. Learners and teachers

could focus merely on the recognition of individual AWL words alone without considering wider and vital aspects of knowing a word including learning and using common collocations and phrases containing these words. It will be of more help if the common collocations and recurrent phrases in a corpus of academic writing can be identified and presented in the writing materials. To offset this shortcoming, concordance lines of the target words can be presented as language data to learners for lexico-grammatical studies.

The workbook developed in Thurstun and Candlin (1997) is an excellent implementation of corpus-informed insights, which offered an innovative way in combining teachers' use of a specialized corpus for materials design and learners trained to analyze the output from the corpus as language researchers. The concordance-based exercises have provided a valuable example for writing teachers to help students develop their academic lexis. This workbook has several merits. As Thompson (2001) commented, the concordance lines of each word were available in the paper format, which allowed them to be immediately accessible. The guiding questions listed in the workbook could also train learners to develop effective corpus analysis skills. The concordance lines had been sifted, so the number of lines was reduced for learners to look through. Nonetheless, one of the shortcomings of this book is that the number of the vocabulary presented was relatively small. There were only 19 words in all. It appeared not to provide sufficient lexical resources for the developing academic writer. If this book was to be used as part of a writing course, learners would need extension activities so that they could explore other related key vocabulary items for each function (Thompson, 2001).

2.2.2. Evaluations of empirical research on students' direct use of corpus

Thurstun and Candlin (1998) suggested that corpus concordance exercises were an effective source input for improving students' vocabulary use in their writing. Webster (2001) advocated a concordance- and genre approach to L2 writing. However, these two studies only gave an account of how to use corpus to help learners develop their language patterns, the students' learning outcomes have not been measured and investigated. In Lee and Swales (2006), the researchers presented an overview of a corpus-informed EAP course aiming at raising students' rhetorical consciousness through exploring the corpora of students' own writing and expert writing. Although this study provided valuable insights into how non-linguistic students could be trained to discover linguistic features and patterns as linguistic researchers, it was based on case studies with four participants.

Yoon and Hirvela (2004) and Yoon (2008) examined how learners evaluated corpus in their writing, and have had shed lights on what issues further research could deal with when learners use corpus as consultation resources in the writing process. The research merely focused on learners' attitudes towards on corpus use. Learners' writing performance has not been investigated and provided as evidence.

As for empirical research focused on the post-writing stage, Chambers and O'Sullivan (2004), Gaskell and Cobb (2004) and O'Sullivan and Chambers (2006) focused on investigating the effects of corpus use on error corrections. In these studies, learners' outcomes were investigated and measured. However, the possible criticism of these research projects is that the students can also correct some of the errors in their writing texts if they are allowed to consult dictionaries or grammars. The research would have been

more solid if two groups were created, one with corpus consultation (Experimental Group), the other with traditional resources, like dictionaries and grammar books (Control Group). Through comparing these two groups' performances on error correction, the findings of the experimental study would be more persuasive.

Studies on students' use of corpus for vocabulary learning (Boulton, 2010; Cobb, 1998; Steven, 1991) have provided statistical evidence that corpus use, especially concordances, helped learners with their vocabulary expansion and gain more lexical knowledge. However, the posttests in these studies were merely designed to measure students' receptive knowledge of the target words. Although the researchers have reported that the experimental group got better score in the vocabulary posttest than the control group did, the ability of using these words as productive vocabulary in their writing has not yet been investigated, and the outcomes of using the target words in their written products have not yet been measured and analyzed.

2.3. Call for further studies

As can be seen above, despite that the literature records a number of empirical studies which attempted to investigate the feasibility and effectiveness of applying corpus at different stages of L2 writing and on vocabulary acquisition, few of them have measured and compared the outcomes of the learners' written products, which makes corpus-informed pedagogy insubstantial (J. Flowerdew, 2002). Most of the previous studies found that data-driven learning was quite applicable and helpful for learners to acquire linguistic knowledge. However, this claim was mainly based on the researchers'

observation and student perceptions of the data-driven learning. None focused on the effects of pre-writing lexical and schemata input on L2 writing with a corpus-informed approach by measuring and comparing the writing outcomes of the control and the experimental groups, thus leaving ample scope for empirical research on this area. This suggests a need to carry out empirical research so as to investigate whether a corpus-informed approach can facilitate learners to obtain collocational patterns of the target items and lexical schemata input for writing, and to what extent it helps them improve lexical use in their written texts.

CHAPTER THREE METHODOLOGY

This chapter describes the current study's methodology including participants, research settings, materials, research design, procedures, research instruments, data collection and data analysis.

3.1. Participants

Forty students participated in the study. They were third-year students majoring in English for Business Purposes, aged between 20 and 22, from the Department of English, Nanhai Campus of South China Normal University. Based on their responses to Questionnaire One (see Appendix 1), these students shared a similar profile. First, they spoke the same L1 language (Chinese). Second, they had a similar duration of English learning (10-11 years). Third, they had the same school timetable, which means the same exposure to English in class.

Their overall English proficiency level was upper intermediate according to the Oxford English Placement Test¹. As for the knowledge on English writing, the participants had learned some basic writing skills in their previous writing classes, such as brainstorming, drafting and revising, organizing an opinion essay and using cohesive devices to make their essays coherent before they took part in the experiment.

The participants were randomly assigned to a control group and an experimental group. Each group had twenty students. A writing pretest conducted before the experiment

¹ http://www.oup.com/shockwave_flash/elt/qpt/qptdemo/

showed there was no statistically significant difference in English writing competence between the two groups. Statistically, 20 students for each group is considered a moderate sample size, which means that the participants are likely to represent a normal distribution of the population or a range of writing abilities that can be expected from a large student population registered in the same course in the same semester.

3.2. Research Setting

The study was implemented in the context of an English writing workshop at Nanhai Campus of South China Normal University. The writing tasks chosen for the study were opinion essay writing, which is a regular and traditional writing task required for university students in Mainland China. There were three tests in total, namely, a pretest, an immediate posttest and a delayed posttest. Each test lasted about 80 minutes. In order to minimize variables, all tests were instructed by the same teacher (the researcher).

3.3. Materials

The materials used for the experimental group were the paper-based corpus-informed concordance exercises. Four factors have been taken into account when developing the concordance exercises: 1) what corpora to be used; 2) how to select the target words; 3) how many concordance lines to be presented to the students and how to sift the lines; 4) how to formulate the concordance queries.

3.3.1. Corpora used

Two types of corpora were used for developing the concordance exercises: a topic-based small corpus and two online free general corpora.

3.3.1.1. Topic-based corpus

The topic-based corpus was composed of texts related to the topics concerning “gambling” and “lottery”, which were obtained from two sources. One was from the online authoritative English news websites; the other was from a ready-made small corpus *LOCNESS* (Louvain Corpus of Native English Essays). One of the advantages of using online resources for text collecting (Nelson, 2009) is:

On the positive side, text collection from the Internet is very fast, and very cheap. Very large amount of data can be downloaded very quickly [...]. [T]he Internet is a good source for certain types of text, especially newspapers and editorials[...]. (p.16)

By adopting the innovative text-collecting approach similar to the one employed in the *ICELite* project (Nelson, 2009), the current researcher, first, identified some authoritative English news websites which contained quality articles on the desired topics. Three websites were selected. They were the *BBC News*, *the Guardian* and *the New York Times*. Once the websites were identified, the researcher searched the key words “gambling” and “lottery” in these three English websites and downloaded the articles related to the subjects. These journalistic articles were included into the corpus as they were on the same topic that the participants were required to write.

The other source was from a sub-corpus of A-level opinion essays on “National lottery” written by British students retrieved from *LOCNESS*. In this corpus, each text has approximately 500-600 words. These essays were not used directly for the current study as they contained a number of grammatical errors and spelling mistakes. Relevant amendments had been made to the texts before they were adopted. All the texts were proofread by an experienced native English teacher of writing, who was also an *IELTS* writing examiner. Twelve articles from *LOCNESS* were selected as they were marked as good writing by the examiner. All spelling mistakes, typos and language errors had been corrected before the texts were used for the study. It is likely to be a criticism that essays written by English native-speaking students may not be a suitable and reliable source for teaching English writing. However, it should also be noted that these revised texts can be deemed appropriate as they deal with the same subject field of the writing task and of the same genre at an appropriate level. They are in fact quite close to the students’ writing compared with the longer academic texts.

3.3.1.2. General Corpora

*COBUILD*² and *COCA*³ (The Corpus of Contemporary American English) were used for retrieving concordance lines of the target words. They were free online resources and easily accessible. *COBUILD* is composed of 56 million words of contemporary written and spoken text. *COCA* has more than 400 million words documented from 1990 to 2009.

² Available at <http://www.collins.co.uk/corpus/CorpusSearch.aspx>

³ Available at <http://www.americancorpus.org/>

3.3.2. *Selecting the target words*

The topic-based corpus described above was used to generate a keyword list for the writing topic with the aid of Wmatrix⁴ through a comparison with the BNC written corpus. The selection of the target vocabulary items was not arbitrary but based on Read's (2004) comments on West's (1953) classic General Service List. As Read (2004) pointed out:

Although the list can be faulted for being dated in some respects [...], [t]he pedagogical value of the General Service List lay in the semantic basis of its selection and presentation. First, words were included not purely on the criterion of frequency but also to achieve efficient coverage of the meanings that learners were most likely to need to express. (p. 149)

In accordance with Read (2004), the target words selected for the current study were based on three criteria: 1) frequency occurring in the topic-based small corpus (each word occurs at least three times); 2) signaling nouns (SNs) often used in opinion essays; 3) lexical schemata for the relevant writing topic (referred to hereinafter as content schemata nouns or CSNs).

Among the chosen items, five words were signaling nouns, frequently used in the argumentative genre. They were *controversy*, *criticism*, *objection*, *situation*, and *effect*. The other five words were content schemata nouns related to the writing topic. Although they were not high-frequency words in the topic-based corpus, they were considered useful for the particular writing topic since they were lexical schemata in nature. For example, words like *jackpot*, *casino*, *charity*, *addiction* and *misery* were considered useful since they were

⁴ An online corpus analysis tool developed by Paul Rayson (2003), available at <http://ucrel.lancs.ac.uk/wmatrix/>

related to the writing topics of “lottery” and “gambling”, and may be used in writing essays on such subjects.

3.3.3. Sifting the concordance lines

First, the concordance lines of each target word were retrieved from a topic-based corpus. The advantage of using the topic-based corpus was that its concordance outputs contained linguistic schemata of the writing topic whereas the disadvantage was the limited number of concordance lines due to the small size of the corpus, which, generally, cannot demonstrate the dominant linguistic patterns of the target words. Hence, in order to get more representative concordance outputs, the general corpora (*COBUILD* and *COCA*) were used as complementary resources, for they contained a substantial number of concordance lines of the searched word. Thus, some concordance lines from the general corpora were selected and added to the concordance output chosen from the topic-based corpus. The number of the lines for each target word was limited to 10-15 only as students might get bored or confused with too many lines.

Not only that the number of the concordance lines was restricted to 10-15, the selection of the lines was also deliberate. The current researcher went through the concordances retrieved from the small topic corpus and *COBUILD/COCA* carefully, selecting the lines that represent the most common collocational features of the target words in accordance with collocational dictionaries (e.g., Oxford Online Collocation Dictionary).

3.3.4. Formulating the queries

Generally, the most common query used in concordance study was to direct learners to study the key word by examining the word either on its left side or on its right side, and find out its linguistic patterning, i.e., what words it often collocates. The query formulated in the pilot study was as follows.

Please study the 5 left words and 5 right words next to *objection*. What kind of words (verbs, adjectives, and etc.) or phrases often follow or are followed by *objection*?

taking money from the lottery 's 'good causes '. The **objection** raised to the lottery of it being 'a tax on the stupid which the lottery money would be broken down. This **objection** was leveled above all at the lottery's 'good causes' Lottery was eventually approved. *Much of the objection* to the National Lottery came from church leaders. harmless fun which won't break the bank. The *main objection to the lottery* is based on the grounds that those who profits to charity, but was rejected. My *personal objection* to Camelot as the lottery organizer is that a large most famous businesses and Families had a *particular objection* to the start of the National Lottery.

All rely on participating viewers who *have no great objection to winning* their moment of glory by doing their best it better to put it another way. The *fundamental objection* to Mother of the Church was that it seemed to place though it has stated that *there is no fundamental objection* and legislation is currently going through the the territory reverts to China in 1997 -- Peking's *main objection* to the airport was that its cost would deplete the

Finally, the *last major objection* was the possibility of one single individual winning a are how bitterly unpopular taxation is; therefore, *objections raised against the lottery* concerned it being marketed areas like horse-racing and casinos. *These objections became much greater* with the introduction of scratch cards week. The proposal of a lottery brought about *many objections and complaints*. There were, and still are, two of conservative government ! *Despite the numerous objections*, the introduction of the lottery has induced a 'fever and psychological drawbacks. *There were two main objections raised against* the introduction of the national lottery holiday or a car. *There were also objections raised to* the amount of money the proposed jackpot was the debate many interested parties stated their *objections to* the National Lottery. One of Britain's most famous

Figure 3.1. Excerpt of the concordance exercises in the pilot study

According to the results of the pilot study, it was interesting to note that most of the participants tended to answer the questions by only noticing the grammatical rule of the target words. A number of students wrote, "The words on the left modifying *objection* are adjectives and verbs, and the word on its right side is a preposition", but the actual adjectives or prepositions collocating with *objection* were not paid attention to by many of them. As a matter of fact, the concordance exercises were not aimed to direct students'

attention to the general grammatical rules of the target words only, but to help students learn their collocational patterns. A more desired answer would be “*objection* can be modified by *strong, personal, main* and *fundamental*; the verb mostly occurred with *objection* is *raise*; we can also *say there is/are objection(s) raised against something*. The preposition following *objection* is *to*”.

There are probably two reasons why the participants missed the point of the activity and tended to only focus on the general grammatical rule of the words: 1) the traditional vocabulary instruction focuses too much on rules rather than on meaning. The traditional EFL class in Mainland China overemphasizes the general grammatical rules in vocabulary learning while the acquisition of formulaic sequences has often been neglected by teachers and students. For example, when they learn *touch*, it is often learned as “*touch* is a transitive verb which can be followed with a noun”, but which nouns to collocate with *touch* seems not very important. This may be part of the reason why Chinese students write “I’ll have an opportunity to *touch the society* after I graduate from the university” due to L1 negative transfer. What they want to express is probably “take a job in the real world” or “gain some real world experience”. 2) The instruction of the query is not explicit and clear enough. The query in the pilot study was a little misleading because it stated “the types of words”, and as a result, students would pay more attention on the word forms instead of the actual words.

Thus, the queries used in the main study were modified according to the feedback given by the learners in the pilot study. Some ideas of Thurstun and Candlin’s design (1997) were employed to ensure that the queries were explicit and clear enough to direct students

to focus on the meaning of the key words, the patterns of language in which they are used and the words that are frequently found before and after these key words.

3.4. Research design

The research was concerned with studying the effects of one particular approach to learning vocabulary for L2 writing by comparing two groups of students (a control group and an experimental group) representing the same population but receiving different treatment. The research design is symbolized as follows:

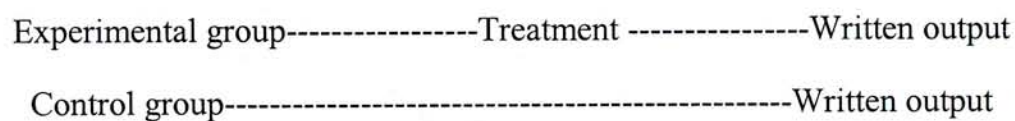


Figure 3.2. Research design

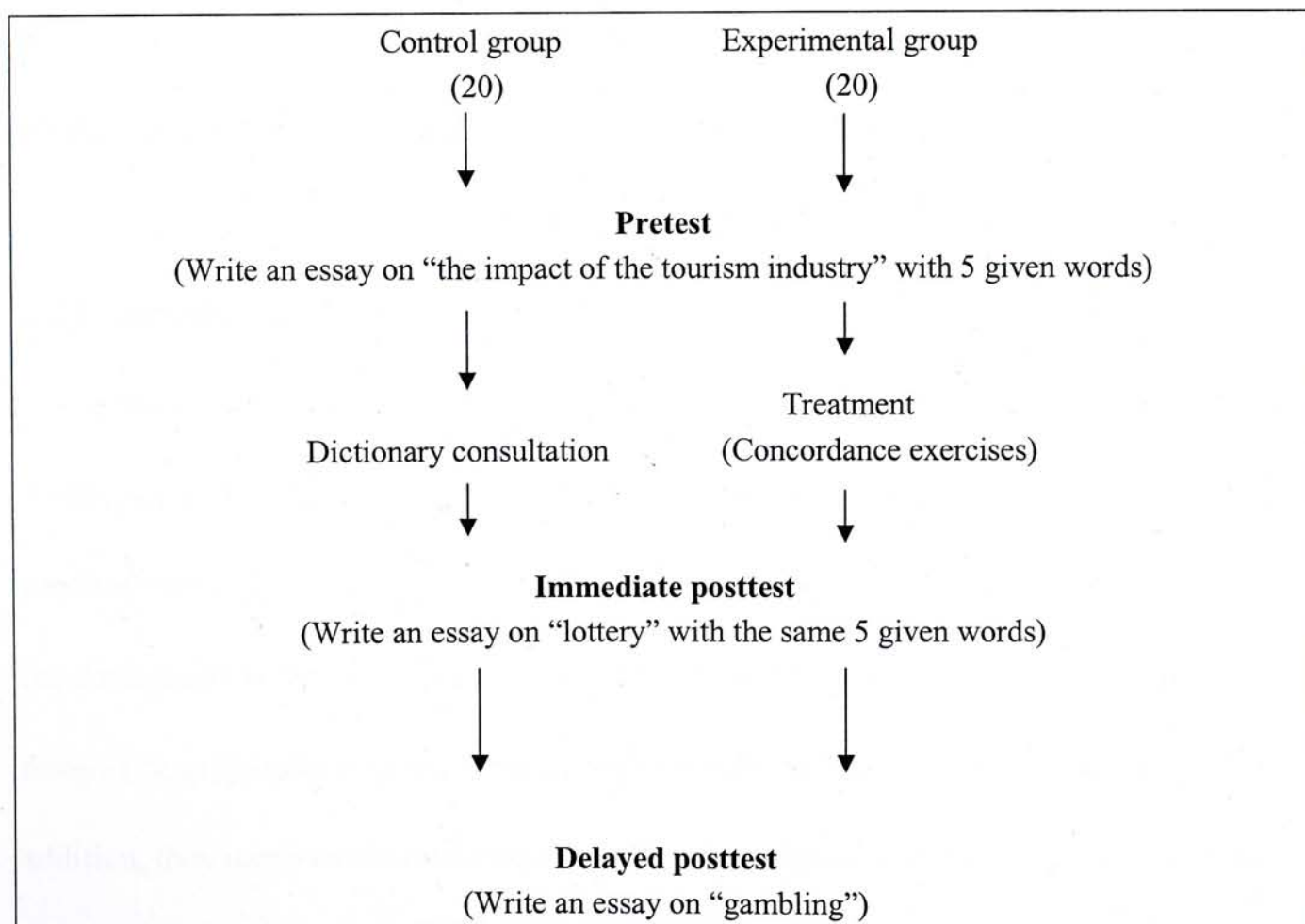


Figure 3.3. Flowchart of the experiment

The two groups were instructed to take three writing tests (a pretest, an immediate posttest and a delayed posttest). Before the immediate posttest, treatment was applied to the experimental group. They were given a set of concordance exercises to learn the target words while the control group did not receive such a treatment. Detailed description of the experiment is illustrated in Figure 3.3.

3.4.1. Pretest

Both the control group and the experimental group were instructed to take a pretest at the outset of the study. Both groups wrote an opinion essay on “the impact of the tourism industry”, and were required to use five given words in their writing. The five words were signaling nouns (*controversy, criticism, objection, situation, and effect*). The purpose of the pretest was twofold: 1) to compare both groups’ overall writing performance; and 2) to obtain data on how the students in both groups used these words before the treatment.

3.4.2. Immediate posttest

Before the immediate posttest, both groups were given a lead-in article related to the writing topic “lottery”. The control group read the article, brainstormed the topic for five minutes with a given list of words (5 SNs and 5 CSNs) which they were allowed to consult the dictionaries if they are unsure of the meaning, and then was directed to write an opinion essay. The experimental group went through the same activities as the control group. In addition, they received the treatment doing the concordance exercises before they started to brainstorm the writing topic. The exercises consisted of 10 sets of concordance lines, each

highlighting the keyword in the center. In other words, 10 words were provided in the KWIC form (Key Word in Context) for these students to study as pre-writing exercises. Among them, five words were the ones that the participants were required to write in the pretest; the other five words were the words related to the writing topic (content schemata nouns). They were *charity, jackpot, casino, addiction, and misery*.

In this writing task, the participants were required to use the same five signaling nouns as in the pretest to write their essays. The use of these five words in the pretest and the immediate posttest was compared between the two groups (inter-subject comparison) and within the experimental group (intra-subject comparison). The purpose was to investigate whether there were changes in using the five signaling nouns in the pretest and the immediate posttest between the control group and the experimental group. It also aimed to examine whether or not, through studying the corpus-informed concordance exercises, learners can use the content schemata nouns properly in their writing.

After the experimental group completed the task, they were instructed to fill in the uptake sheet aiming to elicit which words and phrases in their written tasks were acquired from the concordance exercises.

3.4.3. Delayed posttest

The purpose of the delayed posttest was to examine whether corpus-informed concordance exercises help the students have a longer and better retention of the linguistics knowledge of the target words in the writing task than the control group.

Two weeks after the immediate posttest, the two groups wrote an essay on a topic similar to that of the posttest but with a different title (“gambling”). The purpose of using the similar topic but with a different title was to minimize the practice effect of the writing tasks. In this task, no words were provided in the writing prompt and required to be used in the students’ essays. It attempted to investigate how many words from the concordance exercises were used again in the delayed posttest after a two-week interval and how these words were used in the two groups’ written texts. It was to test whether the experimental group may use more target words in their written texts than the control group. It was assumed that the concordances offered the experimental group an opportunity to discover and formulate the underlying grammatical rules and linguistic patterning.

3.5. Procedures

In this study, the participants in both groups met the researcher four times in total for data collection. The learning venue was a workshop room in the Self-access Center (SAC) at Nanhai Campus of South China Normal University, China. As Table 3.1 shows, in Week 1, a pretest was taken by both groups in the workshop. A questionnaire eliciting background information and perception of writing was distributed to all the students in both groups after the pretest.

In the following week (Week 2), both groups did the immediate posttest, writing an opinion essay on “lottery”. For the control group, traditional writing instruction approach was used in the workshop. To be more specific, the students read a lead-in article related to the writing topic, and then brainstormed before they started to write. The experimental

group followed the same procedures but with the addition of corpus-informed lexical input right after they finished reading the lead-in text. After the experimental group completed the task, they were instructed to fill in the uptake sheet to elicit which words and phrases in their written tasks were learned from the concordance exercises.

Two weeks later (Week 4), both groups did a delayed posttest, writing an opinion essay on “gambling”. Questionnaires on learner’s evaluation of the corpus-informed concordance exercises were administered to the experimental group along with the follow-up interviews in the last week.

Table 3.1. Procedures of the research

Participants	Week 1 Pretest	Week 2 Immediate Posttest	Week 4 Delayed posttest	Week 5 Evaluation on Corpus
Experimental group (20 students)	-Do pretest (write an essay on “the impact of the tourist industry”) (60 mins)	-Read the lead-in article (5 mins) -Study the concordance exercises (15 mins);	-Write an essay on “gambling: (60 mins);	-Do Questionnaire Two on evaluating concordance exercises (20 mins);
	-Do Questionnaire One on background information and perception on writing (20 mins)	-Brainstorm (5 mins) -Write an essay on “lottery” (60 mins) -Fill in the uptake sheets (5 mins);		
Control group (20 students)	Same as above	-Read the lead-in article (5 mins) -Brainstorm (5 mins) -Write an essay (60 mins)	Same as above	Same as above

3.6. Instruments

Three types of instruments were used in the research: questionnaires, learning journals and uptake sheets.

3.6.1. Questionnaires

Two questionnaires were administered to elicit background information and feedback from the participants. They were originally generated in English but translated into Chinese at the time of administering for the sake of better comprehension.

Questionnaire One was distributed to all the participants at the outset of the study to obtain background information on their learning experience in English writing, their overall perceptions of English writing and their English learning strategies.

Questionnaire Two was administered only to the experimental group at the end of the study in order to elicit their feedback and evaluations of the corpus-informed concordance exercises. It consisted of twenty-one items on a Likert scale from 1 to 7 with 1 meaning “strongly disagree”, 2 “disagree”, 3 “Somewhat disagree”, 4 “No opinion”, 5 “somewhat agree”, 6 “agree” and 7 “strongly agree”. These twenty-one items were categorized into three parts: 1) how concordance exercises affected vocabulary learning; 2) how concordance exercises affected L2 writing; 3) what difficulties were encountered when doing the concordance exercises.

3.6.2. Learning journals

Students in the experimental group were instructed to write a learning journal commenting on the concordance exercises and providing suggestions on improving the exercises after the experiment. A list of questions for facilitating students to write their learning journal was distributed to the students by email in order to help elicit students' views and feedback on the concordance exercises (see Appendix Three for the list of questions). Their completed learning journals were sent to the researcher via email as well.

3.6.3. Uptake sheets

Uptake sheets were distributed to all the participants in the experimental group right after they finished the immediate posttest to elicit which words or phrases from the concordance exercises were actually used in their written texts. The data of the uptake sheets provided evidence on whether the concordance exercises helped learners obtain linguistic patterns thus improve their writing quality.

3.7. Data collection and analysis

Three sets of data were analyzed to investigate the effects of prewriting lexico-grammatical input (corpus-informed concordance exercises) on learners' written products: 1) composition data; 2) descriptive data (i.e., questionnaire responses from the participants and uptake sheets); 3) verbal reports (i.e., semi-structured interview). The data analysis was both quantitative and qualitative. SPSS (independent sample *t*-test) was used to assist data analysis of the writing tests.

The main data used for analysis were the participants' written texts produced in the three writing tests, 40 essays for each test, which amounted to 120 essays in total. The analysis of the written products consisted of two components: 1) holistic scoring of the overall writing quality; 2) actual uses of all the target words. All the essays were saved as electronic files in plain text format to form a small corpus. Wordsmith Version 4.0 was used to retrieve the concordance lines of each target word. Copies of all essays were distributed to three external raters for scoring. The essays were also mixed and treated anonymously so that the raters did not know which group each individual essay belonged to.

The procedures of data analysis had three major steps: 1) holistic scoring; 2) analysis of the use of target words in the written texts; and 3) analysis of questionnaire responses and learning journals.

3.7.1. Holistic scoring

By adopting Braine's (2001) scoring approach, three raters marked the essays according to TWE⁵. Each essay was scored twice, each time by a different rater; when the scoring differed by more than one point (e.g., one score of 3 and one of 5), the essays were sent to the third rater to resolve the discrepancy. Pearson's correlation coefficient was computed in order to maintain inter-rater reliability between the two sets of rating by two raters.

⁵ TWE is the abbreviation of TOEFL Test of Written English Guide.

3.7.2. Analysis of the use of target words

A coding scheme was developed to code the data focusing on the use of target words in the written products by both groups. The lexico-grammatical patterns of the five signaling nouns were examined and compared between the control group and the experimental group for inter-subject analysis, and within the experimental group for intra-subject analysis.

To start with, the essays written in the pretest were set up as a small corpus (hereinafter as Corpus 1), those written in the immediate posttest as Corpus 2 and those in the delayed posttest as Corpus 3. Each corpus has two subcorpora, i.e., the control group corpus and the experimental group corpus.

Table 3.2. Students' written corpora

	Corpus 1 (T1: pretest)	Corpus 2 (T2: immediate posttest)	Corpus 3 (T3: delayed posttest)
Subcorpora	Control group	Control group	Control group
	Experimental group	Experimental group	Experimental group

Next, the 20 students in the control group were numbered as c1⁶, c2, c3...c30. Similarly, those in the experimental group were coded as e1, e2, e3...e30. The three writing tests (pretest, immediate posttest and delayed posttest) were marked as t1, t2, t3 correspondently. Accordingly, the usage of each target word by each individual student in each subcorpus was tagged in the following way. Take the tagging of one of the target

⁶ "c" stands for "control group"; "e" stands for "experimental group"; "t" stands for "test".

words *controversy* for example; *controversy* used in the pretest by Student No.1 of the control group was tagged as c1t1, and c1t2 in the immediate posttest. By employing this coding system, the concordances of each target word used by each student in different tests were retrieved with the aid of the corpus analysis tool Wordsmith. The use of these words was analyzed in terms of error-free ratios, error analysis, and grammatical patterns.

The judgments of the usage of the target nouns in the learners' written texts were based on Hymes's (1972, cited from Partington, 1998) adequate theory of language users and language use. Traditionally, the judgments of appropriacy of language use are based on two criteria, "grammaticality" with respect to competence, and "acceptability" with respect to performance. Hymes suggested that the judgments be recognized to be in fact not of two kinds but of four and four questions arise:

- 1) whether (and to what degree) something is formally possible;
- 2) whether (and to what degree) something is feasible in virtue of the means of implementation available;
- 3) whether (and to what degree) something is appropriate (adequate, happy, successful) in relation to a context in which it is used and evaluated;
- 4) whether (and to what degree) something is in fact done, actually performed, and what its doing entails. (p. 281)

In the current research, the appropriacy of the usage of the target words was judged by two native English writing teachers, and further checked against three general corpora, *Word Neighbors*⁷ and *BNC*⁸ and the *COBUILD* corpus.

⁷ An online corpus developed by John Milton at the Hong Kong University of Science and Technology, available at <http://wordneighbors.ust.hk/>. It allows users to look up the contexts of any word within 4 words to the left and 4 words to the right. The total size of the corpus is 141, 000, 000 words.

The inappropriate and less appropriate use of the signaling nouns was further classified into four categories: 1) lexical collocation errors; 2) prepositional errors; 3) word meaning confusion; and 4) word form confusion, based on Flowerdew's (2009) error analysis of signaling nouns. A detailed description of these errors will be elaborated and exemplified in Chapter 4.

The data from the uptake sheets were treated as supplementary evidence indicating which words or collocations the experimental group has learned from the concordance exercises and used as lexical schemata in their writing.

3.7.3. Questionnaire responses

The responses to the Likert-scaled questionnaires in each category were summated and treated as interval data. The mean and standard deviation were calculated by SPSS, and its internal reliability was checked using Cronbach's Alpha. In order to enhance presentation of the questionnaire data, students' responses were coded into three categories, "helpful" "not helpful" and "no opinion," by placing all positive answers (5 "somewhat agree", 6 "agree" and 7 "strongly agree) into the "helpful" category, and all negative answers (1 "strongly disagree", 2 "disagree", 3 "Somewhat disagree") into the "not helpful" category.

⁸ British National Corpus available at <http://www.natcorp.ox.ac.uk/>

3.8. Chapter summary

This chapter describes the methodology applied in the present study. Section 3.1 focused on subjects sampling and grouping along with some demographic information about the participants including their age, major and English proficiency. Section 3.2 briefly introduces the context of the study. Section 3.3 elucidates the criteria and procedures of how the corpus-informed learning materials were developed, which concerned four major elements, i.e., the used corpora (Section 3.3.1), the selection of target words (Section 3.3.2), the sifting of KWIC concordance lines (Section 3.3.3) and formulation of queries (Section 3.3.4). Section 3.4 described the overall research design and the three writing tests used in this study (the pretest, the immediate posttest and the delayed posttest); next followed is Section 3.5 about the research procedures. Section 3.6 describes the instruments, i.e., the two questionnaires, students' learning journals and the uptake sheets. Section 3.7 is concerned with data collections and analysis of three sets of data: holistic scores (Section 3.7.1), analysis of the target words (Section 3.7.2) and analysis of questionnaire responses (Section 3.7.3).

CHAPTER FOUR RESULTS

This chapter describes the major findings in four sections. Section 4.1 compares the holistic scores of the written texts between the control group and the experimental group in the pretest, the immediate posttest and the delayed posttest. Section 4.2 is the major part of this chapter, presenting the qualitative and quantitative analysis of the usages of signalling nouns (SNs) in the three writing tests by the two groups. The analysis of the SNs in students' written output is based on three criteria, i.e., accuracy (Section 4.2.1), complexity (Section 4.2.2) and retention of SNs (Section 4.2.3). Section 4.3 examines the actual use of content schemata nouns (CSNs) by placing a special focus on their frequency and accuracy. Section 4.4 reports students' evaluations of the concordance exercises on the basis of the questionnaires data. The last section (Section 4.5) gives a brief summary of this chapter.

4.1. Holistic scores

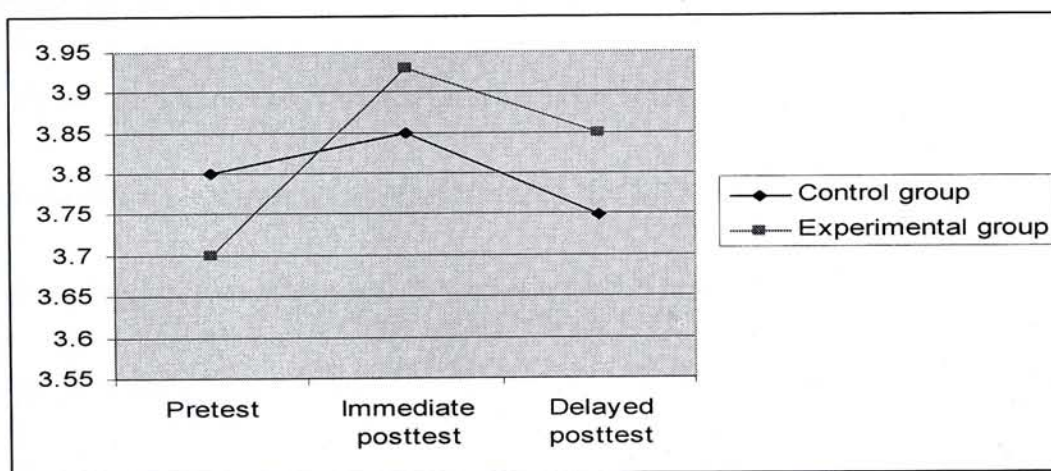
The holistic scores of the pretest, the immediate posttest and the delayed posttest are used to address the research question: Did corpus-informed concordancing help learners improve their overall writing quality? All the papers were marked by two independent raters. If there were any discrepancies between the two scores, the essay was re-marked by a third rater. On the 6-point TWE scale, no paper in the three writing tests scored less than 2 or above 5. The interrater reliability was 0.74. As can be seen from Table 4.1 and Figure 4.1, the mean scores of the control group and the experimental group in the pretest were very close to each other, which are 3.8 and 3.7 respectively. In the immediate posttest,

the mean of the control group increased to 3.85 with a mean improvement of 0.05 while the experimental group's mean score rose to 3.95 with a bigger mean improvement of 0.23. In the delayed posttest, though the two groups' mean scores dropped slightly in comparison with the immediate posttest, the mean of the experimental group remained higher than that of the control group.

Table 4.1. Descriptive statistics of the writing tests

Groups	n	Pretest		Immediate posttest		Delayed posttest	
		Mean	SD	Mean	SD	Mean	SD
Control group	20	3.8	0.49	3.85	0.54	3.75	0.5
Experimental group	20	3.7	0.49	3.93	0.49	3.85	0.48

Graph 4.1. Mean scores of the three writing tests



Based on the independent samples *t*-tests, the control group and the experimental group were not statistically significant different in all three tests ($p > 0.05$).

The paired-samples *t*-test of the pretest and the immediate posttest within the experimental group also shows there was no statistically significant difference between these two tests ($p > 0.05$). The results indicated that the use of corpus-informed concordance exercises had no effect in improving students' overall writing quality in terms of the holistic scores.

However, as writing quality involves a number of variables such as organization, originality of ideas and grammar, it is hard to establish a direct correlation between simply good use of vocabulary and the overall writing quality. In the next, the specific use of the target SNs and CSNs will be examined in detail to show how students acquired these words.

4.2. Use of signaling nouns (SNs)

Analysis of the use of the SNs⁹ was based on the pretest and immediate posttest by examining and comparing the two groups (inter-subject comparison) and within the experimental group (intra-subject comparison) in terms of accuracy, complexity and retention of these nouns. With respect to accuracy, error-free ratios between groups and improved use of the SNs within the experimental group were investigated. In regard to complexity, grammatical structures of the SNs in the immediate posttest were analyzed and compared between groups. Retention of the target items was also measured by comparing the number of occurrences in the delayed posttest between the control group and the experimental group.

4.2.1. Accuracy

Generally speaking, the judgment of language use in a pedagogical context employs a prescriptive perspective whereas the descriptive one is often utilized for linguistic research (Nelson, 2005). The current research tried to adopt a descriptive approach in characterizing

⁹ A signalling noun is the use of an abstract noun the full meaning (realisation) of which can only be made specific by reference to its linguistic context (Flowerdew, 2010). The target SNs in this study are *controversy, objection, criticism, situation* and *effect*.

students' use of SNs. However, due to the educational constraints in which students' written output is often judged as right or wrong in terms of grammatical and semantic appropriacy, the research cannot be entirely free from this prescriptiveness. In view of this, the term "error", though carrying a prescriptive sense, was employed for data analysis in the current research. It should be noted here that the adoption of this term was for maintaining terminology consistence and clear presentation of the findings.

Two English NS (native speaking) teachers were invited to evaluate students' use of the target SNs by categorizing them on a 3-point scale: appropriate, less appropriate and inappropriate (see Table 4.2). If there were disagreements between the two judges, the use of the SN in question was further checked against the two general corpora, *COBUILD* and *Word Neighbors*. If occurrence of such a usage was found in either of these two corpora, the use of the SN concerned was labeled as "appropriate". For other cases where no occurrence (or similar occurrence) of students' usage of the SN concerned was found, the use was labeled as "less appropriate".

Table 4.2. Appropriacy scale

Scale	Category	Description
1	Appropriate	Acceptable to use the target phrase in the context
2	Less appropriate	Grammatically acceptable, but seldom used in argumentative context
3	Inappropriate	Grammatically or semantically incorrect; definitely not used in the context

4.2.1.1. Error-free ratios

For the analysis of error-free ratio, uses of SNs that fell into the two categories of “less appropriate” and “inappropriate” were characterized as errors. Table 4.3 shows the error-free ratios between the control group and the experimental group in the pretest and the immediate posttest. In the pretest, the control group and the experimental group had similar error-free ratios in terms of the use of the five target SNs (39.9% and 37.9% respectively). However, in the immediate posttest, the experimental group’s error-free ratios increased to 88.2% while the control group only to 47.2%. If we take a closer look at Table 4.3, it is easy to discover that the three words *controversy*, *objection* and *criticism* were more problematic with students, as evidenced in the pretest by their lower error-free ratios (below 30%) than those of *situation* and *effect* (above 50%). In the immediate posttest, the error-free ratios of the three problematic SNs in the experimental group increased greatly (*controversy* to 90%, *objection* 88.9% and *criticism* 75%); whereas the control group just made a slight improvement.

Table 4.3. Error-free ratios in the pretest and the immediate posttest

Tests	Words	Control group			Experimental group		
		Occurrences	Errors	Error-free ratio (%)	Occurrences	Errors	Error-free ratio (%)
Pre-test	controversy	16	12	25	19	14	26.3
	objection	18	16	11.1	16	14	12.5
	criticism	18	14	22.2	17	14	17.6
	situation	18	8	55.6	17	8	52.9
	effect	21	3	85.7	30	6	80
	Total	91	53	39.9	99	56	37.9
Immediate posttest	controversy	17	10	41.2	20	2	90
	objection	16	13	18.8	27	3	88.9
	criticism	15	9	40	20	5	75
	situation	20	6	70	22	2	90.9
	effect	23	10	56.5	29	2	86.2
	Total	91	48	47.2	119	14	88.2

4.2.1.2. Types of errors

For better characterizing the errors made by both groups, the errors were classified into four types¹⁰: 1) lexical collocation errors; 2) prepositional errors; 3) word meaning confusion; and 4) word form confusion. Each type of errors was described and exemplified. In each example, inappropriate use of SNs was underlined along with suggested appropriate alternatives presented in brackets at the end of each example.

1) Lexical collocation errors

Lexical collocation is a type of collocation where a lexical item is linked with a lexical word (i.e., noun, verb, adjective). In the current study, the lexical collocation errors mainly concern the use of the SNs with inappropriate or less appropriate lexical verbs. Examples 1, 2 and 3 are instances of this type of errors.

Example 1

A hot controversy goes that tourism helps to develop economy but ruins the local place.

(A hot controversy is that..)

Example 2

In recent years, some controversy on the tourism have aroused because people's unreasonable exploiting the land. (some controversy...has been raised)

Example 3

Some greedy businessmen turned the farmlands into touring places, which gets objection and criticism from the local farmers who depend on the lands for living. (receives some objection..)

¹⁰ The classification of errors in this study is based on Flowerdew's (2009) taxonomy of SN error analysis. A slight modification had been made in order to accommodate the data analysis of this research, i.e., Flowerdew's colligation errors is replaced by prepositional error.

In Example 1, the SN *controversy* was used with the verb *goes*, which was not observed in the general corpora *COBUILD* or *Word Neighbors*. *A controversy goes that + CLAUSE* could possibly result from over-generalizing the existing acquired patterns *A saying/legend goes that + CLAUSE*. In Example 2, *controversy* was used with *arouse*, where the verb *raise* could have been appropriate. In Example 3, the SN *objection* was used with *get*, which sounds a bit awkward and was not seen in either *COBUILD* or *Word Neighbors*.

2) Prepositional errors

This type of error refers to cases where a wrong preposition is used before or after SNs. In Example 4, *controversy* was accompanied by the preposition *of*, where *about* or *over* would have been more suitable. Similar to *controversy*, the SN *objection* seems also quite problematic with students. The wrong preposition *of* was frequently used to follow *objection*, as is illustrated in Example 5.

Example 4

In short, the controversy of the tourism is unlikely to end. (the controversy over..)

Example 5

At last, there is objection of tourism mainly because of ruin bringing from tourists.

(...objection to tourism)

Example 6

It is unlikely that tourism will do harm or have negative effects to the local places or the climate. (have negative effects on..)

Example 6 was a common misuse of the SN *effect*. Students seemed to have difficulty in deciding which preposition, either *to* or *on*, to follow *effect*. In actual language use, *effect* is frequently followed by an *on*-prepositional phrase as complement. When accompanied with a *to*-prepositional phrase, the meaning of the sentence is changed in that the *to*-prepositional phrase functions as adverbial. In Example 6, the preposition following *effect* should be *on*, which would be more semantically appropriate.

3) Word-form confusion

Another type of error was word-form confusion, i.e., misusing SN as adjective or adverb. For instance, in Example 7, 8 and 9, *controversy* was misused as its adjective form *controversial* to modify the noun *topic*. The reason for this error may be due to the overgeneralization of the morphological rule that words ended with the suffix *-y* are often adjectives. Example 10 and 11 are another two instances of noun/adjective confusion in using *controversy*. Another problematic SN is *criticism*, which was often misused as an adjective or adverb. In Example 12, the adjective *critical* would have been more appropriate than *criticism*. Similarly, in Example 13, *through a way of criticism* could have been replaced by *critically*, which apparently sounds more fluent and concise.

Example 7

Nowadays, tourism, one of the world's biggest industries, becomes a controversy topic.
(a controversial topic)

Example 8

Whether we should keep up developing the tourism industry has been a controversy topic.
(a controversial topic)

Example 9

The impact of tourism has become a controversy topic in modern society. (a controversial topic)

Example 10

It is quite controversy that whether tourism is good or poor. (It is quite controversial)

Example 11

People pour into the beautiful scene, problem arises as well, which makes it controversy. (makes it controversial)

Example 12

So, tourism should be taken into consideration in criticism way. (in a critical way)

Example 13

So we should look at it through a way of criticism. (critically)

4) *Word-meaning confusion*

This type of error refers to the misuse of words by confusing its meaning with other words. The existence of such errors may result from unfamiliarity with or misinterpreting the meaning of the target SNs. For instance, in Example 14, the two SNs *objection* and *criticism* were used as words parallel in meaning and connected with a coordinate conjunction *and*. Semantically, such a usage seems inappropriate even though it is grammatically acceptable. These word combinations were not observed in *COBUILD* or *Word Neighbor* either.

Example 14

However, there are still many objections and criticism over the development of lottery.

Example 15

Between objection and agreement, I choose criticism side in all controversy.

Example 15 sounds illogical in that the student writer seemed to confuse the word meaning of *objection* and *criticism*. In this case, either A (*objection*) or B (*agreement*) should have been chosen; however, instead of choosing either A or B, the third factor C (*criticism*) was introduced and chosen in spite of its logical constraint. One of the possible explanations of this error is that the student did not totally understand the meaning and usage of *objection* and *criticism*. To a certain extent, it might be that these two words were probably deemed as synonyms by certain students due to negative L1 interference. In the Chinese language, the meaning of *criticism* (批判 *pi pan*), to a certain degree, has been delexicalized, for some political reasons, by the implementation of *Criticism and Self-criticism Activities* held among the Communist Party members (This is based on the current researcher's own observation. Studies on the semantic change of this word in the Chinese language are worth further investigation for obtaining solid evidence to substantiate the claim). Words collocated with *criticism* gradually assumes a negative semantic prosody. This L1 interference might lead the students to assume that *criticism* could be equivalent to *objection*; whereas its neutral meaning (making judgment and comment) is seldom used or completely ignored.

Table 4.4 shows the frequency of different types of errors made by the experimental group in the pretest and the immediate posttest. It presents an intriguing picture of error distribution, indicating that among these five words, *controversy*, *criticism* and *objection*, are only partially acquired by the students of both groups in view of its high number of errors in the pretest. *effect* and *situation* are comparatively less problematic than *controversy*, *objection* and *criticism*, but there were still some errors, especially

prepositional errors made in the pretest. The students had more difficulties in using the first three words in regard to their lexical collocation and prepositional colligation¹¹. In the pretest, word meaning confusion is the most frequent error (n=17). Most of these errors (10 occurrences) resulted from the misuse of *objection* and *criticism* together as synonyms which are then connected using the coordinating conjunction *and*.

Table 4.4. Errors distributions of the experimental group

Tests	Words	Lexical collocation errors	Prepositional errors	Word-meaning confusion	Word-form confusion
Pretest	controversy	3	3	1	7
	objection	6	3	5	0
	criticism	1	1	7	5
	situation	0	2	4	0
	effect	0	2	1	0
	Total	10	11	17	12
Immediate posttest	controversy	0	0	1	0
	objection	1	1	1	0
	criticism	2	0	1	2
	situation	1	0	0	0
	effect	0	2	0	0
	Total	4	3	3	2

In comparison with the pretest, four types of errors decreased in number in the immediate posttest, particularly word-meaning confusion errors. It can be seen here that the use of concordance exercises proved to be quite effective, as is evidenced by the number of correct use of target SNs. Through studying the collocational patterns of target SNs, students were equipped with more ways of using the SNs in their interlanguage system. So far as the two SNs *objection* and *criticism* are concerned, students were more resourceful in

¹¹ Colligation was defined as the collocation of a lexical and a grammatical item (Partington, 1998)

using these two nouns instead of placing these two words together using a coordinating conjunction *and* as found in the pretest. The number of lexical collocation errors of *objection* reduced dramatically in the immediate posttest (n=1) as opposed to the pretest (n=6). No prepositional errors of *controversy*, *criticism* and *situation* were found in the immediate posttest.

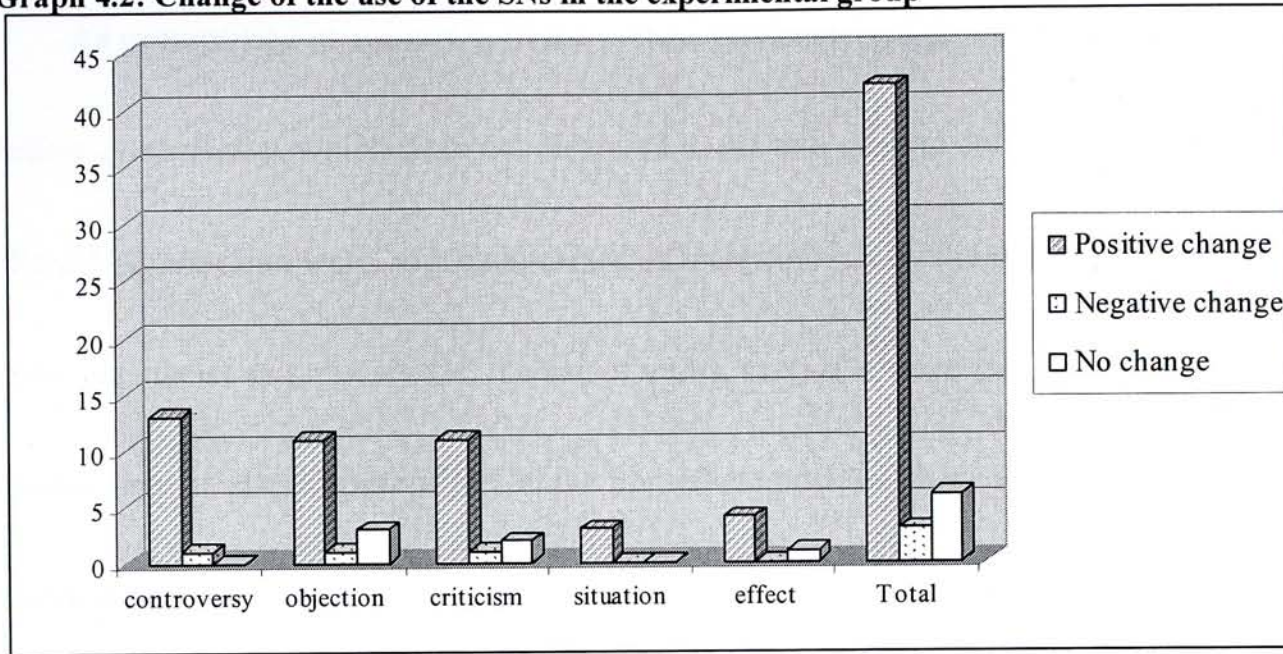
4.2.1.3. Improved use of SNs in the immediate posttest as opposed to the pretest

In the above, the use of SNs in the pretest and the immediate pretest among the experimental group has been characterized and classified into three major categories: appropriate, less appropriate and inappropriate. However, whether students in the experimental group made improvement in terms of the correct use of the target SNs was not touched upon. For addressing this issue, uses of these SNs were studied and grouped into three major categories based on the comparison of their usage in the pretest and the immediate posttest. The three categories were termed as: positive change, negative change and no change. The descriptions of the three types of change are illustrated below (see Table 4.5).

Table 4.5. Description of SNs usage change between the pretest and the immediate posttest

Change	Description
Positive change	Inappropriate or less appropriate use of SNs in the pretest but appropriate use in the immediate posttest
Negative change	Appropriate use of SNs in the pretest but less appropriate or inappropriate in the immediate posttest
No change	Inappropriate or less appropriate use of SNs both in the pretest and the immediate posttest

Graph 4.2: Change of the use of the SNs in the experimental group



As can be seen from Graph 4.2, the instances of positive change (42 instances in total) outnumbered negative change (3 instances) and no change (6 instances). Let's take a further look at the positive change of the three SNs *controversy*, *objection* and *criticism*, which are illustrated in Table 4.6, Table 4.7 and Table 4.8 respectively.

Table 4.6. Positive change of *controversy* in the experimental group.

Student	Pretest	Immediate Posttest
S26	However, while tourism is developing, the <i>controversies</i> of it also arise among the people.	However, there has also been some <i>controversy</i> over the lottery games.
S32	But when so many people pour into the beautiful scene, problem arises as well, which makes it <i>controversy</i> .	However, lottery games have provoked considerable <i>controversy</i> .
S19	As tourism has become one of the world's biggest industries, a <i>controversy</i> topic , whether the influence of tourism is good or not has also been raised.	People's awareness of the spending of lottery funds has caused considerable <i>controversy</i> .
S13	In short, the <i>controversy</i> of the tourism is unlikely to end.	People waste a lot of money and time on it, which stirs considerable <i>controversy</i> .
S10	A hot <i>controversy</i> goes that tourism helps to develop economy but ruins the local places and climate at the same time.	As welfare lottery is playing an increasingly bigger role in China since 13 years ago, it has provoked a <i>controversy</i> surrounding it.

As exemplified in Table 4.6, in the pretest, students used the prepositions *of* and *on* to follow *controversy*, which were not observed in the two general corpora *COBUILD* and *Word Neighbor* and were both categorized as inappropriate by the two NS raters. Apart from the misuse of prepositions, the use of verbs, such as *arouse* and *go*, as predicates of *controversy*, is also inappropriate. In the immediate posttest, these students made improvement in using *controversy* by collocating it with a more appropriate preposition *over* instead of *on* and *of*. More appropriate verbs, such as *cause*, *stir* and *provoke*, were also used as predicates of *controversy*.

Table 4.7. Positive change of *objection* in the experimental group

Student	Pretest	Immediate Posttest
S2	There are more and more people who are more in <i>objection of</i> tourism and the impact of tourism is in controversy.	The miseries brought about by lottery lead to great <i>objection raised against</i> lottery.
S18	Nowadays, some people have an opinion of <i>objection</i> towards tourism.	There are more and more people raising <i>objection to</i> the lottery.
S33	Many people are for the tourism while others hold the idea of <i>objection</i> .	Much of the <i>objection to</i> the lottery is about the spending of the collected money.
S1	At last, people hold the <i>objection of</i> tourism mainly because of ruin bringing from tourists.	The main <i>objection to</i> the welfare lottery is that people waste a lot of money and time on it.

As indicated in Table 4.7, *objection* was only used as Object (in terms of sentence element) with one type of syntactic pattern (i.e., *the Agent has objection to something*) in the pretest. Apart from a limited number of syntactic variations, the lexical collocations of

objection produced by the students sound rather improper and redundant, e.g., *have an opinion of objection* (S18) and *hold the idea of objection* (S33). The word *objection* itself already denotes a kind of opinion or viewpoint; thus examples produced by S18 and S33 are rather wordy. Another inappropriate use is the preposition *of* to follow *objection*.

However, in the immediate posttest, improvement has been made in terms of more variety of syntactic patterns and correct collocations. Take the sentences produced by S33 and S1 as examples; *objection* was used by both students as *Theme* in a declarative sentence by adopting a syntactic structure *Subject (objection) + Copular verb BE + Subject*

Complement. An appropriate verb *raise* was also employed to collocate with *objection* in the immediate posttest.

Table 4.8. Positive changes of *criticism* in the experimental group

Student	Pretest	Immediate Posttest
S5	We'd better think of it with <i>criticism</i> .	Although welfare lottery has significant effects on the economy, it also faces <i>criticism over</i> its negative effects.
S26	However, we should understand the situation in <i>criticism</i> .	Another <i>criticism is that</i> they are games which are easy to get addicted to.
S12	Some people think tourism is quite good as it benefit people a lot, while some <i>criticism against</i> it because tourism may do harm to our environment.	The <i>criticisms can be rejected</i> because the case (mentioned) above rarely happens.
S6	Between objection and agreement , I choose <i>criticism side</i> in all controversy.	Some people praise that lottery will be a boost to the country's domestic consumption while the lottery also came under intense <i>criticism</i> by other people.

Let us have a look at the positive change of the word *criticism*. In Table 4.8, S5 and S26 used prepositions *with* and *in* to precede *criticism* in the pretest, although the word *critically* could have been used as a more appropriate candidate. The sentence (i.e., ..., *while some criticism against it because...*) produced by S12 in the pretest is grammatically incorrect in that a verb is missing in the main clause. It can be improved by using *There + BE structure* (e.g., *while there is criticism against it*). Similarly, the sentence (i.e., ..., *I choose criticism side in all controversy*) produced by S6 in the pretest is grammatically correct but semantically inappropriate. These two students, after studying concordance-based exercises, made improvements accordingly. In the immediate posttest, it can be seen that two students (S5 and S12) turned to a correct use of appropriate verbs, i.e., *reject* and *face*, to predicate *criticism*, and one student (S6) used a prefabricate unit (e.g., *something came under intense criticism*).

Although the occurrences of positive change far outnumbered the other two categories, cases of negative change and no change in the immediate posttest still existed. These cases are illustrated in Table 4.9 which gives a list of examples of no change and negative change. For example, the error made by S3 in the pretest (i.e., ...*in criticism way*) was repeated in the immediate posttest (i.e., *hold a criticism problem*), where *criticism* was used to modify abstract nouns *way* and *problem*. It might be that this particular student intended to say that *the development of tourism should be taken into consideration critically*. Similarly, in the immediate posttest, she repeated the wrong pattern by using *criticism* as a premodifier to follow a noun, while the correct use should probably be *we should look at this problem critically*.

Table 4.9. No changes in the experimental group

Student	Pretest	Immediate Posttest
S3	So, tourism should be taken into consideration in criticism way.	We should hold a criticism problem.
S26	Though it causes many problems and makes some bad effects on our environment, in objection , it is good for the development of our world.	The main objection of it is based on the fact that the rate of winning the prizes is far too small.
S15	Furthermore, from tourism, the countries can show every aspect to the world and have a better effect on foreigners.	Many people hold the objection against them because they ignore the bad effect of lottery games especially to the young children who have less control of themselves than adults do.

The output by S26 is quite noteworthy. In the pretest, the error may be attributed to word meaning confusion by mixing up the meaning of *objection* with *contrast*. In the immediate posttest, the student attempted to use *objection* in a more acceptable syntactic construction *SN (as Subject) + BE + Subject complement*, which is more appropriate in terms of grammatical structure than the sentence used in the pretest. The only error made here is the preposition *of*, which should be *to*. Compared with the pretest, the use of the SN *objection* by S26 in the immediate posttest was improved syntactically and also clearer in meaning. Judging from the fact that the student committed a global error in the pretest while a local error in the immediate posttest using this word, it may, to a certain extent, be concluded that he has made improvement even though the attempt is not fully successful.

4.2.2. Complexity

Complexity in the current research concerns the variety of grammatical patterns of the five SNs. Throughout this research, patterns were defined based on the notions used in corpus-based linguistic analysis of English grammar. An encompassing definition of the term “grammatical patterns” can refer to the patterns of natural language both lexically (collocations, recurrent word combinations) and grammatically or syntactically. These grammatical patterns range from simply lexico-grammatical patterns (e.g., verb complement) to complex syntactic patterns (e.g., zero relative clauses). (Meunier, 2002, p.121). Huston and Francis (2000) defined a pattern as:

[A] phraseology frequently associated with a word, particularly in terms of the prepositions, groups of [words], and clauses that follow the [target] word. Patterns and lexis are mutually dependent, in that each pattern occurs with a restricted set of lexical items, and each lexical item occurs with a restricted set of patterns. (p.3)

Coxhead and Byre (2007) defined lexico-grammatical patterns as “frequently occurring combinations of words and grammar, where a particular word generally requires particular grammar” (p. 130). That is, the verb *required* can be followed either by an infinitive or by a that-clause. However, the most commonly used combination involves *required* followed by an infinitive. The combination of *required* and the infinitive is a lexico-grammatical pattern.

In the current research, grammatical patterns were used as a terminology to characterize the actual uses of the target SNs in the students’ written texts. The patterns included word combinations (lexico-grammatical patterns) of the target SNs and the syntactic structures of

these words. These grammatical patterns were classified into different categories from a functional perspective. Appropriate uses, together with inappropriate and less appropriate ones were all incorporated for analysis. The purpose of including inappropriate and less appropriate uses for investigation was to make a balanced comparison. Since the error-free ratios of the control group were much lower than those of the experimental group and there were relatively fewer occurrences of appropriate uses by the control group, the comparison would appear uneven if only appropriate uses were included. Apart from this, it is also beneficial to investigate the inappropriate uses of the SNs because this type of written output is the developmental phenomenon of interlanguage.

Since noun phrases (NPs) headed by SNs were quite common in students' written texts, it is also worth examining the patterns to see whether there are any differences between the control and the experimental groups. Thus, all NPs headed by the five SNs (including appropriate, less appropriate and inappropriate uses) were within the scope of the current investigation.

4.2.2.1. Grammatical patterns

It is interesting to find that four SNs (*controversy, objection, criticism and situation*) have more syntactic variations in terms of their sentence patterns used in the experimental group than those of the control group. Detailed information on the patterns of these four SNs is illustrated in the following tables.

Table 4.10. Grammatical patterns of *controversy* in the immediate posttest

Functions	Grammatical structures	Control group	Experimental group
		Patterns	Patterns
Object	V + SN (active voice)	1. It has caused some <i>controversy</i> .	1. Lottery game has provoked <i>controversy</i> . 2. The negative aspects of lottery games has caused <i>considerable controversy</i> 3. Lottery games are facing <i>any controversies</i> . 4. Lottery has stirred <i>many controversies</i> . 5. *We cannot deny <i>the numerous controversies</i>
Subject	SN + V (passive voice)	2.*A <i>controversy</i> is put into heat discussion.	6. <i>Controversies</i> have been triggered again 7. <i>The numerous controversies</i> are mostly tied to both legal and moral issues. 8. Controversy over the lottery has been raised since it was born... 9. A controversy over lottery is stirred .
Subject complement	Copular verb BE + SN	3. This is a <i>controversy</i> .	----
Notional subject	There BE + SN	4. There is <i>controversy</i> on welfare lottery	10. There has been some <i>controversy</i> .

* Referring to the inappropriate and less appropriate use of the SNs.

In Table 4.10, the most common grammatical pattern found in the control group was *controversy* used as *Subject complement* in grammatical structure *There + copular verb BE + a controversy*. This pattern was used by about two thirds of the students. Lexical verbs, such as *provoke*, *face* and *stir*, were not found in this group. Apart from this, few students in the control group attempted to use the target noun in passive voice. There was only one attempted use of passive voice (i.e., *A controversy is put into heat discussion*), which was unsuccessful because the student used an inappropriate verb *put* to modify *controversy*. In direct contrast, the experimental group was more resourceful in using the SN *controversy*,

as was evidenced by the variety of lexical verbs (e.g., *provokes, cause, face, stir, trigger, raise*) used to collocate with *controversy*. In addition, they also made improvement in terms of grammatical accuracy. To a large extent, the experimental group outperformed the control group in both areas of syntactic variations and correct grammar use.

Table 4.11. Grammatical patterns of *objection* in the immediate posttest

Function	Grammatical structures	Control group	Experimental group
		Grammatical Patterns	Grammatical Patterns
Object	V + SN (active voice)	1. *It has gained <i>objection</i> because it...	1. Some projects have received considerable <i>objections</i> .
		2. *Our government gives <i>objection</i> to gambling.	2. It has brought about many <i>objections</i> and complaints
		3. Some may still maintain an <i>objection</i> on welfare	3. Some people may raise an <i>objection</i> that
		4. I want to present my <i>objection</i> to lottery	4. I put forward a personal <i>objection</i> to... 5. Many people hold the <i>objection</i> against lottery games
Subject	SN + V (passive voice)	5. *An <i>objection</i> should not be given to welfare lottery any more.	6. Many <i>objections</i> were raised against lottery
	SN + Copular Verb BE + <i>that</i> -clause	-----	7. Another <i>objection</i> is that the chance to win the price is too slim.
Notional subject	SN + Copular Verb BE + PP	-----	8. Much of the <i>objection</i> to the lottery is about the spending of money.
	There BE + SN	6. * There should not be any <i>objection</i> on it.	9. There are still many <i>objections</i> to lottery games.

As for the use of *objection*, both groups demonstrated more syntactic varieties as compared with the SN *controversy* (see Table 4.11). Similar to the SN *controversy*, the

experimental group again outperformed the control group in terms of syntactic varieties and correct grammar use.

As is indicated in Table 4.11, the usages of SNs by the experimental group were found in three subcategories, i.e., *objection* used as (1) *Object*; (2) *Subject*; and (3) *Notional subject*. For instance, in the subcategory where *objection* functions as subject, the control group only had one pattern (i.e., SN + *V*), while the experimental group had three patterns (i.e., (1) SN + *V*; (2) SN + *Copular Verb BE + that-clause*; and (3) SN + *Copular Verb BE + PP*). This indicates that students, through studying the concordance-based exercises, had activated their interlanguage system by raising their awareness of the grammatical patterns with which the SN occurs, which enabled them to reproduce more sentence patterns in their own writing. The other significant difference is reflected in the variety of lexical verbs used. The control group had some attempted use of lexical verbs as predicator of *objection*, such as *give objection to* and *gain objection*. However, their uses of grammatical patterns are not appropriate. The experimental group, on the other hand, used more appropriate lexical verbs (e.g., *receive*, *bring about*, *raise*, *hold*). Clearly, such usages were acquired from the concordance-based exercises. To a certain degree, the concordance-based exercises seemed rather effective in improving the correct use of lexical verbs to collocate with the SNs.

Table 4.12 shows the grammatical patterns of *criticism* produced by the control and the experimental groups. As noted in the table, the featuring differences in using *criticism* between the two groups are reflected in two areas: (1) the SN + *V* pattern in passive voice; (2) the *Prep.* + SN pattern. As for the SN + *V* pattern, there was no attempted use of *criticism* in passive voice by the control group, while in the experimental group, the

students were able to use lexical verbs, such as *reject* and *raise* to collocate with *criticism* in passive voice, which to some extent adds more objectivity to the tone (e.g., *The criticism can be rejected; Strong criticisms are raised against...*). The other noticeable distinction is the *Prep. + SN* pattern. After observing the concordance lines of *criticism*, the experimental group used the newly acquired pattern *under + attributive adj. + criticism* in the immediate posttest, however, this prefabricated unit was not found in the control group. The results, again, show that students in the experimental group were more resourceful in using the SNs in view of their syntactic variety.

Table 4.12. Grammatical patterns of *criticism* in the immediate posttest

Function	Grammatical structures	Control group	Experimental group
		Patterns	Patterns
Object	V + SN (active voice)	1. *Others give much objection and <i>criticism</i> .	1. *I put forward my <i>criticism</i> over the opinion. 2. It also faces <i>criticism</i> over its negative effects.
Subject	SN + V (passive voice)	----- -----	3. The <i>criticism</i> can be rejected because the case seldom happens. 4. Strong <i>criticisms</i> are raised against gambling.
	SN + copular verb BE + <i>that</i> -clause	2. Another <i>criticism</i> on lottery is that it increases the unstable factors of the society.	-----
Notional subject	There BE + SN	3. There are both <i>criticism</i> and praise.	5. There have been great <i>criticisms</i> of the lottery from all walks of life.
Complement of PP	Prep. + SN	4. Welfare lottery is full of <i>criticism</i> .	6. The lottery also came under intense <i>criticism</i> . 7. Though under great <i>criticism</i> , lottery game survive its way by playing an important role of benefiting the economy.
			8. *Some people are immediate in their <i>criticism</i> of the lottery.

Table 4.13. Grammatical patterns of *situation* in the immediate posttest

Function	Grammatical structures	Control group	Experimental group
		Patterns	Patterns
Object	V + SN	-----	1. This will worsen <i>the situation</i> further. 2. It may exacerbate <i>the situation of poverty</i> . 3. To make <i>the situation worse</i> , people who have won the lottery would have the risk of being kidnapped by bad guys.
Subject	SN + V	1. <i>The situation will become</i> more serious if we do not take any action.	4. Some bad <i>situation could happen</i> as well.
Complement of PP (PP as Adverbial)	Prep. + SN	2. <i>In this situation</i> , welfare lottery is no more a good thing but an evil. 3. * <i>On this situation</i> ,... 4. <i>Under this situation</i>	5. In this <i>situation</i> , it is very harmful to the society. 6. Under <i>this situation</i> ,
Complement of PP (PP as Object complement)	Prep. + SN	-----	7. Lottery will leave some people in <i>situation of risk</i> . 8. These people will throw themselves into a desperate <i>situation</i> . 9. If they don't make fortune, it will trap people in the <i>situation of poverty</i> . 10. People in desperate <i>situation</i> tend to turn to the lottery

Now let's move to the grammatical patterns of *situation* illustrated in Table 4.13. The most noticeable difference is the variety of grammatical patterns. Apparently, the experimental group produced more patterns than the control group. The control group used *situation* in merely two patterns, i.e., SN + V (e.g., *The situation will become more serious*) and Prep + SN (*situation* embedded in a PP which functions as Adverbial in the whole

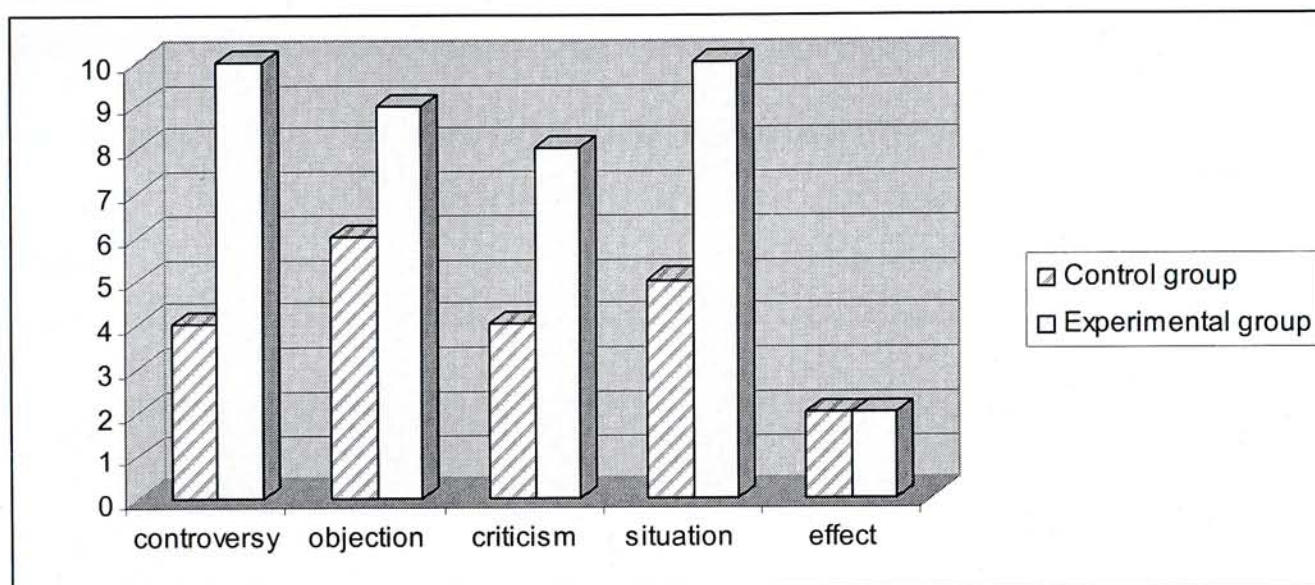
sentence, e.g., *In this situation, welfare lottery is no more a good thing*), whereas the experimental group used two additional patterns: *V + SN* (e.g., *worsen the situation; exacerbate the situation*) and *Prep + SN* (*PP* functions as *Object complement*, e.g., *Lottery will leave some people in situation of risk*).

Compared with the experimental group, the number of patterns produced by the control group appeared to be rather limited. It is worth noting that in relation to accuracy, there were not so many differences in the error-free ratios between the control group and the experimental group in the immediate posttest (see Table 4.3); however, if we take a further look at the complexity of the patterns, it is surprising to find that the experimental group outperformed the control group in terms of grammatical patterns. This finding reveals that corpus is beneficial to learners not only because it is helpful in improving learners' accurate use of the target SNs, but also in assisting them to acquire new sentence patterns. In contrast, most of traditional references, such as dictionaries, seem not so effective in enhancing learners' awareness of sentence patterns. The KWIC (Key Word in Context) form of concordance-based materials, in this regard, has the advantages of attracting learners' attention to the form of language use.

Graph 4.3 compares the number of grammatical patterns of the five SNs in the immediate posttest between the control group and the experimental group. Overall, the four SNs, except *effect*, were used with more grammatical patterns by the experimental group than the control group. The grammatical patterns of *effect* were equal in both groups. They fell into two types: (1) *V + effect* by collocating *cause* and *have*; and (2) *effect + Copular*

Verb BE as Subject (e.g., *Another negative effect is that lottery games have caused many crimes*).

Graph 4.3. Distribution of grammatical patterns in the immediate posttest

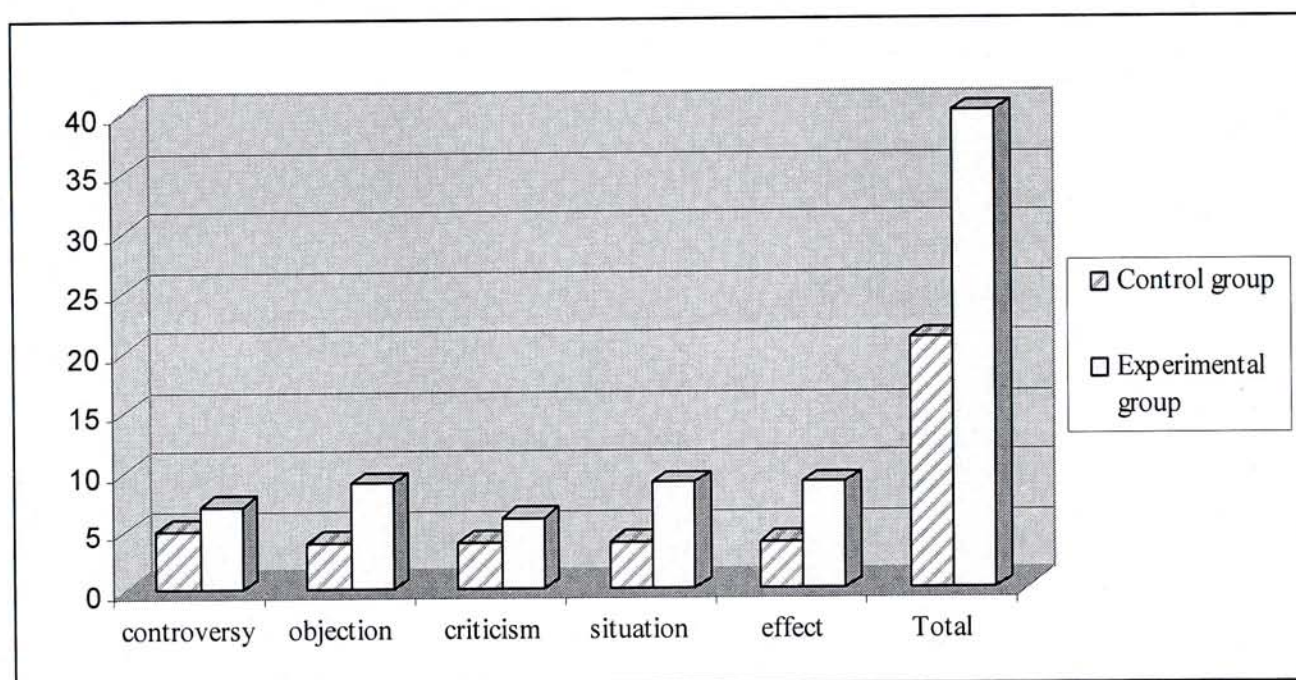


4.2.2.2. NPs headed by SN

Another pattern under investigation was the NP patterns. NP headed by SN was classified into two types: 1) Premodifier + SN; 2) SN + postmodifier. The overall comparison of NP patterns between the two groups was quantified in Graph 4.4 and Graph 4.5. On the whole, in terms of frequency, the experimental group had more variety of NP patterns than the control group. 39 occurrences of Premodifier + SN and 15 of SN + Postmodifier were found in the experimental group while only 20 and 11 in the control group. As can be seen from Table 4.14, the control group tended to use determiners as premodifier to collocate with the SNs, such as possessive articles *my* and *their*, definite article *the*, indefinite article *a* and *an*, and referential pronoun *this*. The adjectives modifying the SNs were also limited in numbers, such as *bad*, *good*, *heated* and *great*. The patterns in the experimental group presented a more intriguing picture. The most salient

difference from the control group was the use of adjectives. There were more variety of adjectives used to modify the target nouns, such as, *numerous* and *enormous* to precede *controversy*, *strong* and *considerable* to quantify *objection*, *intense* and *main* to modify *criticism*, *desperate* and *complicated* to describe *situation*, and *harmful* and *significant* to modify *effect*. These adjectives were less informal than those in the control group (*good*, *bad* and *great*). They were more appropriate in argumentative essays.

Graph 4.4. Distribution of *Premodifier* +SN in the immediate posttest between the two groups



Graph 4.5. Distribution of SN + *Postmodifier* in the immediate posttest between the two groups

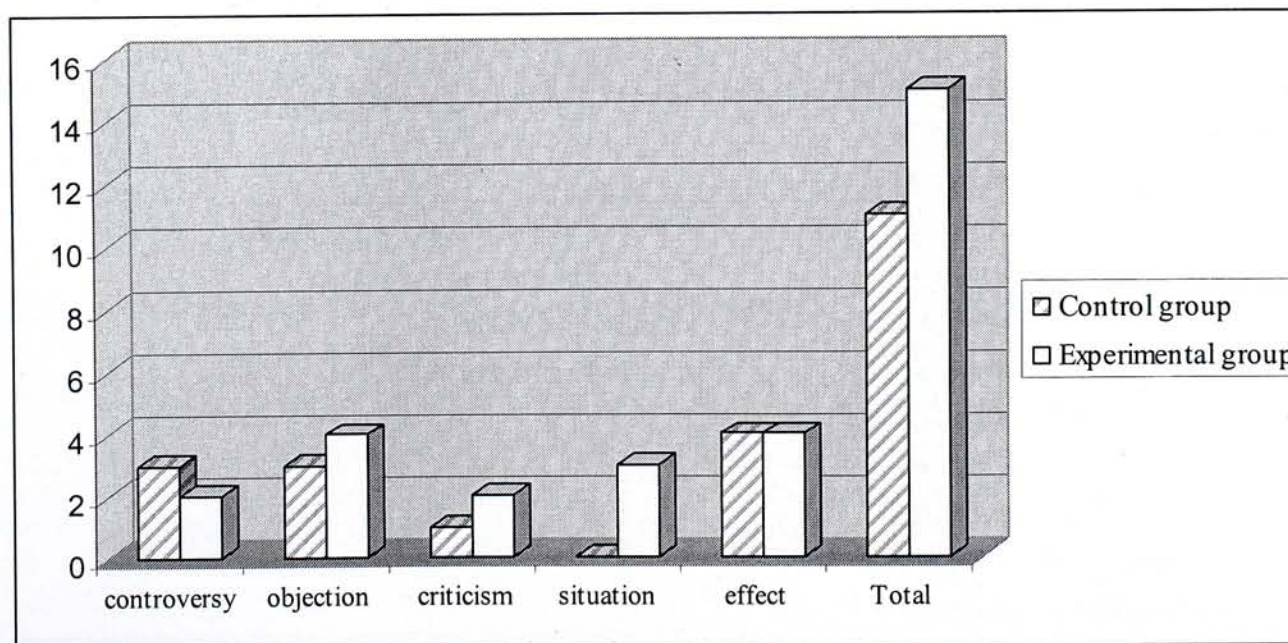


Table 4.14. Patterns of *Premodifier* + *SN* in the immediate posttest between the two groups

Structure	SN	Control group	Experimental group
		Patterns	Patterns
Premodifier + SN	controversy	1. a controversy	1. numerous controversies
		2. much controversy	2. considerable controversy
		3. some controversy	3. intense controversy
		4. the heated controversy	4. strong controversies
		5. a big controversy	5. enormous controversies
			6. some controversy
			7. many controversies
	objection	1. their objections	1. considerable objections
		2. my objections	2. a particular objection
		3. an objection	3. great objection
		4. any objection	4. strong objection
			5. the main objection
			6. my personal objection
			7. many objections
			8. another objection
			9. a lot of objection
	criticism	1. the criticism	1. great criticisms
		2. this criticism	2. strong criticism
		3. another criticism	3. intense criticism
		4. a lot of criticism	4. the social criticism
			5. serious criticism
			6. another criticism
	situation	1. this situation	1. a negative situation
		2. the situation	2. a desperate situation
		3. a bad situation	3. the harmful situation
		4. the real situation	4. a serious situation
			5. the general situation
			6. this complicated situation
			7. some special situation
			8. every situation
			9. some situation
	effect	1. the bad effect	1. a bad effect
		2. a good effect	2. its harmful effects
		3. many big effects	3. a terrible effect
		4. greater effect	4. the serious effect
			5. a significant effect
			6. another negative effect
			7. the positive effect
			8. the potential harmful effects
			9. numerous harmful effects

Table 4.15. Patterns of SN + Postmodifier in the immediate posttest between the two groups

		Control group	Experimental group
Structure	SN	Patterns	Patterns
SN + PP	controversy	1. * <i>controversy on</i> the welfare lottery	1. <i>controversy over</i> the lottery
SN+V-ing		2. * <i>controversy of</i> the effect	2. <i>controversies surrounding</i> lottery games
		3. * <i>controversy in</i> this issue	
SN + PP	objection	1. <i>objection to</i> gambling	1. <i>objections to</i> the view mentioned above
		2. * <i>objection on</i> welfare lottery	2. <i>objection against</i> lottery
		3. * <i>objection of</i> welfare lottery	3. * <i>objection of</i> it
SN + relative clause			4. <i>objection raised</i> against
SN + PP	criticism	1. * <i>criticism on</i> welfare lottery	1. <i>criticism of</i> the lottery
			2. <i>criticism over</i> its negative effects
SN + PP	situation	---	1. the <i>situation of</i> poverty
			2. (be in) <i>situation of</i> risk
			3. (be in) <i>situation of</i> risking investing more money
SN + PP	effect	1. <i>effect on</i> the social development	1. <i>effect on</i> (the society)
		2. <i>effect of</i> welfare lottery	2. <i>effects of</i> (the lottery games)
		3. * <i>effect to</i> the people	3. * <i>effect to</i> the quality of people's life
SN + zero relative clause		4. <i>effects made</i> by welfare lottery	4. <i>effects caused</i> by gambling

Table 4.15 shows the postmodifier + SN patterns used by the control group and experimental group in the immediate posttest. Both groups used all the patterns despite their disproportionate distribution among different SNs. The most common postmodifiers were the prepositional phrases (PP) and the zero relative clauses. Both groups attempted to use these two patterns to follow the SNs. As can be seen, the experimental group demonstrates more patterns in using the postmodifiers. For example, the usage of

surrounding to follow *controversies* was provided in the concordance exercises and a number of students seemed to notice this pattern and it was used in their own written output. Clearly, the usage *controversies surrounding lottery games* seems much more proper as a sentence element. Also, it can be noted here that there are more inappropriate prepositions used in the control group than in the experimental group. Based on the results, it can be seen that the experimental group outperformed the control group in regards to the variety and appropriacy of postmodifiers used to follow SNs.

4.2.3. Retention of the target patterns

Table 4.16. Occurrences and errors of the target words in the delayed posttest

Test	Words	Control group		Experimental Group	
		Occurrences	Errors	Occurrences	Errors
Delayed posttest	controversy	3	1	12	2
	objection	0	-	9	1
	criticism	0	-	6	0
	situation	1	0	13	0
	effect	6	0	14	0
	Total	10	1	54	3

Table 4.16 shows the occurrences of the five target signaling nouns between the two groups in the delayed posttest (the free writing task). 54 occurrences of the target SNs were found in the experimental group while only 10 in the control group.

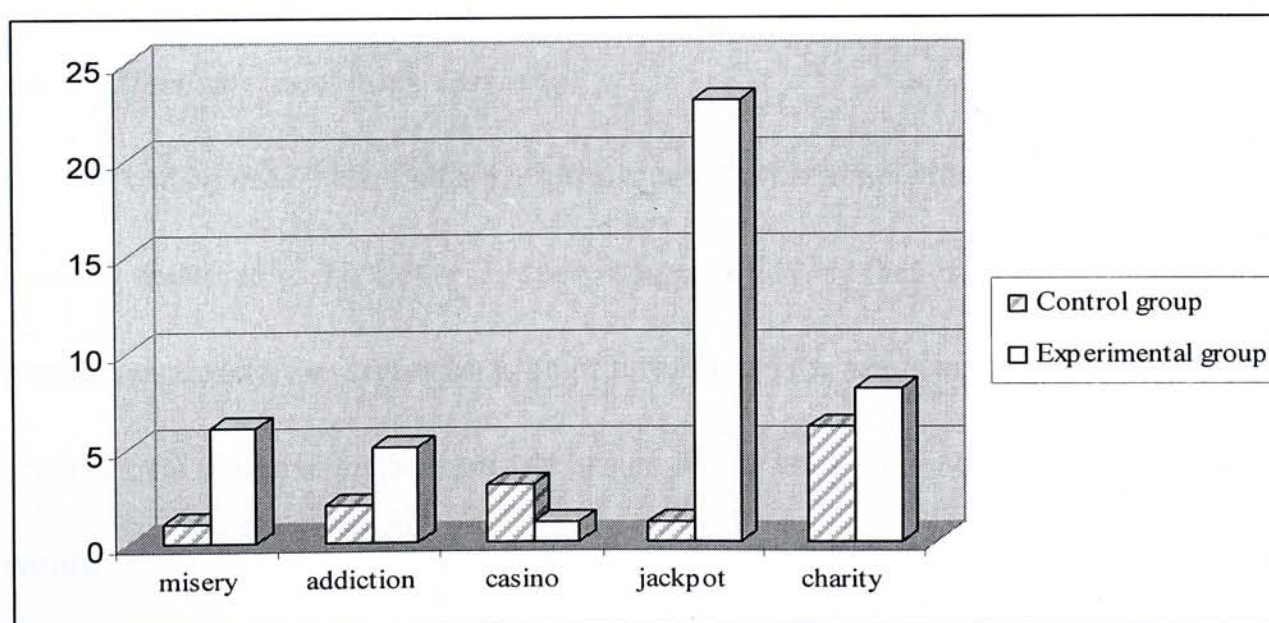
In other words, the occurrences of the target words in the experimental group outnumbered the control group by more than 5 times. Moreover, students in the experimental group were quite accurate in retaining the lexico-grammatical patterns of these target words in view of the fewer occurrences of errors. The errors made in the

delayed posttest by the experimental group were mainly prepositional errors, such as S24 who wrote *the controversy of the legality of gambling has been provoked in China*, and S15 *I have objection of gambling*. The writers repeated the errors made in the pretest.

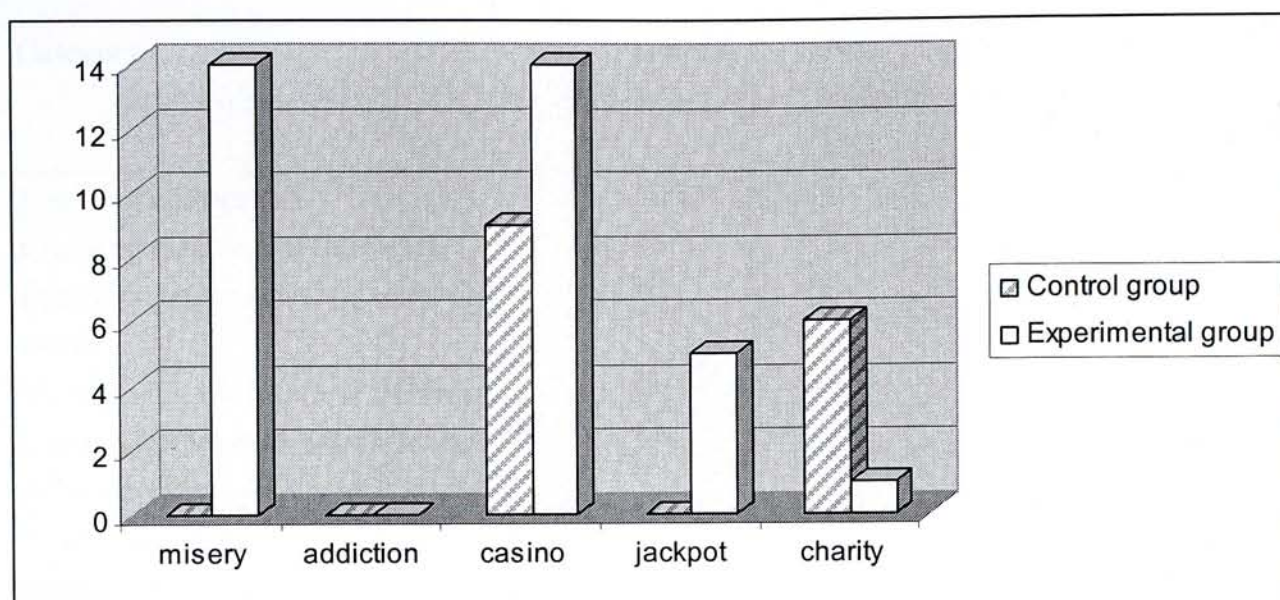
4.3. Content Schemata nouns

As indicated by Graph 4.7 and 4.8, the occurrences of the target content schemata nouns in the experimental group are more than the control group, particularly the frequency of *jackpot*, which was a new word to most of the students. Likewise, the delayed posttest reflected the same picture as that in the immediate posttest, the words related to the writing subject *gambling*, such as *casino* and *misery* are more frequently used in the experimental group than the control group.

Graph 4.6. Distribution of content schemata nouns in the immediate posttest



Graph 4.7. Distribution of content schemata nouns in the delayed posttest



4.4. Evaluation of the concordance exercises

The survey results were the findings of the questionnaires following up the experiment. The results of the questionnaires were placed into four main categories: 1) the effects of concordance exercises on vocabulary learning; 2) effects on L2 writing; 3) difficulties in doing the concordance exercises and 4) overall evaluation.

4.4.1. Effects on vocabulary learning

As can be seen from Table 4.17, the mean scores of students' reaction to vocabulary learning clustered in the 5.50-6.30 score range, indicating that, overall, the majority of the students evaluated the exercises as helpful resources for acquiring collocational and grammatical patterns, memorizing the usage of the words and incidentally learning new words.

Table 4.17. Effects on vocabulary learning (n=20)

Category	Helpful (%)	Not Helpful (%)	No opinion (%)	Mean*	S.D.
Learning collocation	100	0	0	6.30	0.65
Learning meaning of the words	95	5	0	5.80	0.89
Learning grammatical patterns of the words	95	5	0	5.90	0.96
Memorizing the usage of the words	90	10	0	5.50	1.05
Learning other new words incidentally from the exercises	95	5	0	5.90	0.91

*1: very unhelpful, 2: unhelpful, 3: somewhat unhelpful, 4: no opinion, 5: somewhat helpful, 6: helpful, 7: very helpful.

It is worth noting that the average score of 6.30 regarding collocation learning stood out from the other categories. Unlike the other four categories in which students' opinions were varied, almost all the students' responses in this category were positive about the concordance exercises and deemed that they helped them acquire the collocation patterns of the target words. One student commented:

Using the collocations [learned] from the vocabulary exercise also enables me to have a rich expression. Take *controversy* as an example, I used to use "have controversy" and this seemed to be dull. However, through the vocabulary exercise, now I know more usages like "provoke/reignite controversy", which could polish my writing better.

In addition to improving lexical collocation, many students concurred that concordance-based exercises helped them with their prepositional colligations.

To be honest, most the words offered are known to me, but I merely have vague ideas of them, such as just the Chinese meanings. Actually I was not sure about the exact prepositions that

follow the words. Take controversy for example, I always use “of” as its preposition and seldom have I used “about” or “over”. [As a] matter of fact, the “of” is inappropriate to well express the meaning of reasons for controversy while only either “about” or “over” can do. So the exercises can help us to know more about the words and help us to use prepositions in a correct way.

It also raised the student awareness of collocations in vocabulary learning. Student

LHY wrote:

After doing the exercise, I know the common usage of the word and the phrases the word often collates. In this way, I can have a good command of these words and my essay would be academic if I know the collocation. In the future reading and writing, I will pay more attention to the collocation[s].

In contrast, with regards to memorizing the usage of the words, about 90% of the students thought that such a type of exercises was somewhat helpful and 10% believed that it was unhelpful. The reason why some students considered it not useful for memorizing the usage of the words can be explained away by the reports in the learning journal. Student

Zhong wrote:

As far as I am concerned, it did help me in some way, but the help is little. The vocabulary exercises is really a good way to improve our writing skill because it shows how to use a word in different ways and help us to know more about the word even though some words we have been familiar with. However, it does not help us very much. To learn a lot of different usage of words is a good thing, but to remember all of these usage[s] in a short time is really a challenge to me, let alone put to use [in writing]at once.

Another Student Wen remarked, “I think it helps not so much because I cannot memorize them when I want to use it in my writing. I need to refer [to] the exercise if I intend to use the new collocations.” Similar to Student Wen, Student L commented:

I always forget some collocations of the words in the writing... I know it's my personal problem, nothing to do with the exercises. But for me it is not so helpful because I always have difficulties to memorize them.

4.4.2. Effect on L2 writing

Table 4.18 provides an intriguing look at the students' evaluations on the specific aspects in which the exercises helped their L2 writing.

Table 4.18. Effects on L2 writing (n=20)

Category	Helpful (%)	Not Helpful (%)	No Opinion (%)	Mean*	S.D.
Learning expressions related to the topic	95	5	0	5.40	0.75
Obtaining ideas for their writing	80	15	5	5.30	1.38
Useful resource for learning vocabulary for English writing	100	0	0	5.85	0.58
Help memorize the key words related to the writing topic better	85	10	5	5.15	1.03
Improving overall writing quality	95	5	0	5.20	0.83

*1: very unhelpful, 2: unhelpful, 3: somewhat unhelpful, 4: no opinion, 5: somewhat helpful, 6: helpful, 7: very helpful.

Similar to vocabulary learning, the attitudes towards concordance exercises for L2 writing was quite positive. The mean responses clustered from 5.15 to 5.85. However, several results are worth noting. First, compared with vocabulary learning, the mean scores

of student perceptions on L2 writing were, on the whole, relatively low. Only one category (useful resource for learning vocabulary) had an upper-5 score (5.85) while the rest were less than 5.50. This may indicate that, in general, students perceived the corpus-informed activity as a more effective means for vocabulary learning than enhancing writing performance. For instance, with respect to obtaining ideas for writing, the mean score was only 5.30, and the individual score given by the students varied largely clustering from 2 to 7 on the scale.

Among them, 15% of the students had a negative response (one gave a 2-score and two gave a 3-score); whereas, on the other end of the 7-point scale, 15% believed it very helpful, giving a 7 score to this category. Why the evaluation differed so strikingly may be explained by the data from the learning journal. For those who held unfavorable views to corpus use for gaining ideas, the examples demonstrated in the concordance lines were too many. The overwhelming examples may become a distraction when organizing their ideas, and adversely affected their writing performance. For these students, they had a problem in organizing the provided information rather than not having adequate ideas. As Student Wu commented in the questionnaire:

The vocabulary exercises helped with my writing task on “lottery” in two aspects. For one thing, I was able to familiarize myself with the proper usages of the key words by reading the listed examples so that I wouldn’t use the words in a wrong way. For another, the examples gave me some ideas about the topic. However, there are so many examples that I almost lost my patience with them. Besides, I may not be able to learn them by heart for future use. Personally, I would prefer useful phrases and simple sentences.

This problem was also reflected in the students' written output as one examiner pointed out:

The students' essays are just like magpies, putting all the examples and ideas together without any logic and unity. They seem to have little capacity to sort the good from the bad or to present a cohesive position.

This finding is a subtle reminder that topic-specific concordances are likely to have a negative effect on students' writing quality due to information overflow provided in the concordance output. Presumably, lack of ideas is one of the difficulties often encountered by students in L2 writing. It is assumed that topic-based concordance lines can provide students with some content schemas helpful for improving ideas in their writing. However, the results reminds us that the concordance lines need to be used with caution because, according to the abovementioned reports, not all the students seemed to be able to benefit from this type of exercises. The reasons that some students thought corpus concordances unhelpful for L2 writing is further explored and presented in 4.4.3.

4.4.3. Difficulties in doing the concordance exercises

Table 4.19. Problems in doing the concordance exercises (n=20)

Category	Agree (%)	Disagree (%)	No Opinion (%)	Mean	S.D.
Time-consuming	75	20	5	4.50	1.31
Unfamiliar vocabulary	80	10	10	4.75	1.16
Cut-off sentences	80	20	0	4.85	1.63
Too many sentences	65	30	5	4.40	1.42
Limited number of sentences	30	60	10	3.45	1.19
Difficulty in formulating the overall rules of the usage of the words	45	50	5	4	1.68

*1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: no opinion, 5: somewhat agree, 6: agree, 7: strongly agree.

It is interesting to find that, although the mean responses in Table 4.17 and Table 4.18 indicate that the majority of the students had a favorable attitude towards corpus use for vocabulary learning and L2 writing, Table 4.19 reveals a different intriguing look. 75% of responses showed that it was time-consuming to do this form of exercises; despite the fact that half of the students reported they did not have difficulty in formulating the overall rules of the usage of the words.

As the Student LJR wrote:

I think there are too [many] contents which cost our lots of time. It would be better if there is less exercise or we just [under]line the answer in the content and not need to write it out.

About 80% of the students thought that they had difficulties in doing the concordance exercises due to the cut-off sentences and the new words in concordances. The cut-off sentences hampered them to fully understand the information presented in the concordance output. They remarked that the cut-off sentences prevented them from understanding the examples or the viewpoints from the sample texts in the concordance lines as Student YLL commented:

However, the examples of the words using are not so perfect because some of them are just a part of a sentence, and we don't know what the whole meanings of the examples are. So these examples can not well express the exact using of the words. Some of them just show the verbs or prepositions that can be used with them.

The cut-off sentences are one of the main reasons why some students thought that topic-based concordance lines failed to provide them with ideas related to the writing topic. The students commented that it would be of help for them to grasp the ideas more thoroughly if they could go further into the full context of the target word, read the complete sentences or even the whole paragraph. This may also explain why their overall writing quality had no statistically significant difference.

4.4.4. Overall evaluation

Table 4.20. Overall evaluation on the concordance exercises

Category	Agree (%)	Disagree (%)	No Opinion (%)	Mean	S.D.
Overall evaluation on vocabulary learning	95	5	0	5.90	0.91
Overall evaluation on L2 writing	95	5	0	5.50	0.94
Increasing confidence about using the words in L2 writing	95	5	0	5.70	0.92
Expecting more concordance exercises in future writing	85	15	0	5.50	1.27
Prefer learning words from the concordance lines than being taught by teacher	70	30	0	4.90	1.51

*1: strongly disagree, 2: disagree, 3: somewhat disagree, 4: no opinion, 5: somewhat agree, 6: agree, 7: strongly agree.

As noted in Table 4.20, vocabulary learning again ranked the top among the categories. About 95% of the students were favorable to concordance exercises, reporting it was helpful for vocabulary learning and increasing confidence about using the words in L2 writing. However, it is worth noting that only 70% of the students preferred learning the words from concordance than being taught by the teacher, given that the responses to this question were neutral with a mean score of merely 4.90.

Most of the students had a conservative attitude to this form of exercises, either agree or disagree. This can be furthered supported by the evidence found in the learning journal.

Student ZJW wrote:

I think it helps not so much...If the sentences can be one by one to explain the usage of the words by the teachers, then give the examples and give a space for us to imitate or make a sentence that can help our memory.

What is most worth noting is students' belief in DDL (data-driven learning) and TDL (teacher-driven learning). On the whole, the students were relatively conservative about DDL given that its mean score falls in the "no opinion" category. Despite of a neutral response to DDL, 85% of the students expected to have more corpus exercises in the future.

S34 is a representative that evaluated the concordance exercises negatively. It seemed her attitudes towards the exercises correlates to her performance of using the target nouns in writing. She responded to the preference of DDL with a score of 2 (disagree) and 3 (somewhat disagree) for expecting more practice in the future. As to the difficulties in doing the exercises, S34 agreed that she had problems when formulating the linguistic patterns due to the cut-off sentences. Her unfavourable response to the corpus activity influences her performance in using the target nouns in her writing. Take the use of *controversy* for example; she used the word inappropriate in the pretest, and no positive changes in the immediate posttest and delayed posttest.

4.5. Chapter summary

The current chapter has reported the findings of this controlled experiment in details. It consists of four sections: 1) holistic scores; 2) use of the signaling nouns (SNs); 3) use of the content schemata nouns (CSNs); and 4) students' evaluations.

In Section 4.1, although there was no statistically significant difference in the overall writing quality of the three writing tests between the control group and the experimental group, the mean improvements of the experimental group were consistently higher than those of the control group in the immediate posttest and the delayed posttest.

Section 4.2 compared the actual use of five SNs in students' written output between the two groups regarding two linguistic aspects, 1) accuracy and 2) complexity. Based on the error-free ratios and occurrences of grammatical patterns of the SNs, the experimental group apparently outperformed the control group by using the SNs more accurately and producing more variety of grammatical patterns in their written texts. It was also found in Section 4.3 that the experimental group intended to use more CSNs than the control group did in the immediate test and the delayed posttest.

Students' evaluations in Section 4.4 provided further insights into the role of corpus-informed concordance exercises on students' L2 writing. On the whole, a majority of the students had positive attitudes towards this form of lexical input at the prewriting stage. Concordance exercises were ranked as most effective for improving vocabulary use, in particular collocations and colligations. With regard to obtaining ideas for writing and assisting memorization of target patterns, students' responses varied greatly with a relatively low mean score as compared to vocabulary use.

CHAPTER FIVE DISCUSSION

This chapter interprets and explains the major findings of this study by revisiting related theories in Second Language Acquisition (SLA) and comparing the findings with previous research. The effects of corpus-informed concordances on L2 writing are discussed in accordance with the research questions posed earlier in Chapter One. Section 5.1 restates the holistic assessment of students' written output and evaluates the impact of corpus-informed concordances on L2 overall writing quality by examining the influence of content schemata nouns (CSNs) input. Section 5.2 focuses on the micro-linguistic level, elaborating on the ways corpus-informed concordances improved lexico-grammatical patterns of the five target signalling nouns (SNs) in students' written output. Section 5.3 evaluates students' response to corpus use in L2 writing. Section 5.4 attempts to propose a tentative model of corpus-informed L2 writing instruction based on the findings of the current research and the frameworks developed by other researchers in previous research.

5.1. Did the corpus-informed approach improve students' overall writing quality?

By comparing the holistic scores of the pretest, the immediate posttest and the delayed posttest, it was found that there was no statistically significant difference in overall writing quality between the control group and the experimental group. This finding indicated that the corpus-informed concordances did not have a significant impact on students' writing outcomes in terms of overall writing quality. To a certain extent, this might suggest that studying SNs and CSNs from the KWIC (Key Word in Context) concordance learning

materials was of little help for improving L2 learners' overall writing performance. It should be noted, though, that L2 writing performance is subject to a number of linguistic variables besides vocabulary use. General ideas, text structures, grammatical use, and coherence and cohesion are equally vital factors that determine the overall writing quality. As an experimental study, this research was set up to elicit results within a fixed period of time. Its effects on overall writing quality might not be as salient since the subjects were instructed to study a limited number of nouns (5 SNs and 5 CSNs only) by doing concordance exercises. It is quite possible that the input of these 10 target items would not make a difference in holistic writing scores within such a short period of time (4 weeks in total). In order to attain more desirable results of the effects of topic-specific corpus on overall writing ability, a long-term study with more lexical input in the form of KWIC concordances is worth carrying out.

In addition to the limited number of lexical input, the input of content schemata nouns also influenced the overall writing quality. Five content schemata nouns (CSNs) were given to the control group and the experimental group at the prewriting stage for brainstorming. Through studying the corpus-informed concordance exercises at the prewriting stage, it was hypothesized that learners can obtain lexical schemata related to the writing topic through studying the provided topic-related concordance lines, thus helping them improve their overall writing quality. Based on the findings, the hypothesis was not supported. Although the experimental group used more target CSNs in the immediate posttest and the delayed posttest than the control group, the input of CSNs did not have a desirable effect on helping students gain related ideas and background

knowledge immediate to the writing subject. In fact, it turned out to undermine students' writing performance, particularly in relation to the organization of content. In probing into the reasons why concordance exercises failed to improve students' writing, it is important to refer back to students' own evaluations as they provided some insider information. Two major causes are attributable for the negative impacts of the CSNs on L2 writing: 1) cut-off sentences; 2) culture-loaded information in concordance lines.

5.1.1. Cut-off sentences

In regards to the effects of concordances on L2 writing, most of the participants concurred that they had difficulties in obtaining the ideas from the concordances due to the truncated sentences. What they could get from the concordance lines was only some incomplete information which seemed to be of little help for content development. To make it worse, they were overwhelmed with the incoherent and fragmented lines. The overflow of information adversely hampered students' organization of content in their writing. This problem reported by the participants was in line with the comments of the two independent raters that the essays of the experimental group contained many viewpoints but were rather inadequate in terms of coherence. These responses suggested that the input of content schemata in the form of cut-off sentences is of little help for improving writing content. To some extent, specialized corpora provide students a short-cut to read a set of topic-based and genre-related texts. Concordances can operate as "condensed reading" (Gabrielatos, 2005, p.8), a bypass for students to obtain content schemata instead of having extensive reading in order to obtain their vocabulary related to writing topics. With the aid

of corpus analysis software, keywords of a topic-specific corpus with same genres can be generated easily. Through exposure to cut-off concordance lines, learners can analyze the linguistic forms at the surface level; however, to further understand the contextual information, they need to get access to a full context. In other words, for the purpose of developing ideas in writing, more contextual use of the target words or phrases must be available and accessible so that students can observe them with a fuller context, thus reducing their difficulty of interpreting meaning and usage of the target words due to the cut-off sentences.

5.1.2. Culture-loaded information in concordance lines

Another aspect worth mentioning is the culture-specific characteristics of concordance lines. The small topic-based corpus consisted of essays written by British students, news reports from *BBC*, *Guardian* and *New York Times*. These resources contained some concepts and background information that the participants might not be familiar with due to their particular cultural background. For example, one of the CSNs *charity* is a relatively high-frequency word in the small corpus compiled for the current research (54 occurrences in a two-thousand-word corpus), indicating that it was quite relevant to the writing subject. However, for the participants, it seemed that they had some difficulties in connecting this word with the writing topic mainly due to the Chinese context in which, *charity*, is barely related to the subjects of *gambling* and *lottery*. In this case, students may have lacked the necessary schemata to connect the two subjects. This cultural difference may explain why

it is hard for students to understand the connection of *charity* with *lottery* and *gambling*.

This echoes Adam Kilgarriff¹²'s claim:

[C]oncordances are hard to read. Whether they are fragments of sentences or full sentences, they come without a discourse context and they are likely to throw at the learner a whole array of complexities and difficulties...[O]ften, culture-specific inferences need to be made to understand the sentence, making matters very difficult for learners of the language. In order to learn anything about word meaning from the corpus line, it is first necessary to decode the line itself: for learners, this is no trivial task.

To make it worse, the cut-off sentences made the cultural-loaded lines more difficult to comprehend. Some students suggested that if they could access the larger context of the SCNs and read the full context, they would have more ideas about why and how the target items were used in the articles. As Sabine (2007) pointed out, in a pedagogical context, learners can benefit from a small pedagogically mediated corpus by adopting “a discourse-based approach, which focuses on the analysis of the individual texts in the corpus, and of linguistic means of expression in relation to their communicative (situational) and cultural embedding” (Sabine, 2007, p.1). Under such arrangement, each individual text can be accessed by learners in order to further analyze the meaning of the target items under investigation.

We may draw a conclusion here that the input of CSNs does not automatically enrich students' linguistic knowledge relevant to the writing subject and enhance writing quality.

¹² Source available at http://scholar.google.com.hk/scholar?cluster=10992087205563452707&hl=en&as_sdt=2000&as_vis1

Its application to writing instruction should be taken into account with caution. More activities, such as group discussions about organizing the content schemata at brainstorming stage, need to be incorporated in the classroom. Truncated concordance lines, on the one hand, might reduce the information load, allowing a focus on linguistic form in short and multiple contexts, showing various usages simultaneously and without distraction of longer stretches of discourse (Cobb, 1999). On the other hand, for obtaining meaning, it seems necessary to present a more extended context to learners since the cut-off sentences may hamper full comprehension of the target word. For learning the linguistic forms, in particular lexico-grammatical patterns of a lexical item, perhaps the number of words in a four-line concordance is roughly equivalent to that in one full-sentence context, as Stevens (1991) has claimed. The multiple contexts provided in concordances are a more efficient way for learners to acquire linguistic patterns than just being exposed to a single longer context (Levy, 1990). However, for a focus-on-meaning study, a more extended context rather than the truncated lines are necessarily desirable.

5.2. Did the corpus-informed approach improve vocabulary use in students' writing?

Although there was no significant difference in the overall writing outcomes between the control and the experimental groups, the mean improvement of the experimental group in the immediate posttest was much higher than that of the control group. This could suggest that corpus-informed concordances did exert certain positive effects on students' writing quality in the short term.

One of the possibilities that the experimental group had a better mean improvement in the immediate posttest in comparison to the control group could be attributed to more accurate and complicated structural use of the target SNs after the treatment, i.e., concordance-based lexical input. The quantitative and qualitative analysis of the usage of target SNs in both groups' written output indicated that corpus-informed concordances are especially effective in helping learners to obtain lexico-grammatical patterns of the target SNs when compared with dictionary consultation, which confirmed the findings reported in previous empirical research (e.g., Boulton, 2009; Koo, 2007). This type of learning resources has a notable effect on improving language use in L2 writing, particularly in relation to collocational uses and grammatical patterns. The effect is indeed twofold, which concerns both acquisition and production. First, it could help learners acquire a variety of collocational patterns. Second, acquisition of collocational patterns in turn enables learners to generate more sophisticated and proper syntactic patterns. In the following sections, the discussions will elaborate on the possible reasons that concordance exercises help improve students' acquisition and production of these syntactic patterns.

5.2.1. Interface of lexis and syntax

First, let's take a look at the finding of accuracy of the target SNs. The relatively low error-free ratios of the target SNs in the control and experimental groups in the pretest, to some extent, revealed that rule-based learning may not be adequate for improving learner language use in L2 writing, at least for the students at upper-intermediate level. As a matter of fact, before this experiment, both groups had already acquired the basic meaning of the

target SNs. They were able to recognize these words in reading and listening. Also, they were familiar with the grammatical rules of a noun, being aware that in a sentence, a noun can be preceded by a verb, modified by an adjective or followed by prepositional phrases. However, this prior knowledge does not automatically lead to successful use of the target SNs in writing. When the learners attempted to use these words as productive vocabulary, most of them encountered problems in selecting appropriate words to collocate with the SNs. Lack of lexico-grammatical patterning led to an overgeneralization of acquired routines or patterns, resulting in a number of language errors. Let's take *objection* and *criticism* for examples. In the students' written texts, these two SNs were found to be used together in *N + and + N* structure regardless of their semantic restrictions. Also, the students tended to assume that SNs could be predicated by any lexical verbs in a *V + N* structure. With this assumption, they ended up choosing inappropriate lexical verbs to collocate with the SNs (e.g., *gain objections* or *a controversy is put into hot discussions*).

Another linguistic feature of sentence patterns of the SNs in the control group was *simplification of grammatical structures*. The students appeared to use a limited number of syntactic structures and failed to produce complex structures. The common structures used by the control group are *There BE + SN* and topic closing structure *Theme + Copular verb BE + SN* (e.g., *This is a controversy*). Some students did attempt to use lexical verbs to predicate the SNs; nevertheless, most of these attempted uses turned out to be unsuccessful by collocating with inappropriate verbs. The three instances below exemplify the students' difficulty of selecting an appropriate verb.

Example 1:

Let's take the criticism, find out the bad causes of tourism and try to make it more useful.

Example 2:

Our government gives objection to it.

Example 3:

Other people show their objections and arise some criticism.

In contrast to the control group, the qualitative analysis of the experimental group's written texts in the two posttests revealed that after the treatment (observation of corpus-informed concordances), the experimental group was able to use a variety of lexical verbs to fill the *V* slot in *V + SN* pattern, such as, *trigger*, *provoke* and *stir up* to collocate with *controversy*; *face* and *reject* with *criticism*; *brought about* with *objection*, *worsen* and *exacerbate* with *situation*. It is interesting to find that the acquisition of these lexical collocations, in return, assisted the students to generate the grammatical structure *Lexical verb + SN* not only in active voice but also in passive voice. Conversely, it is striking to find that very few instances of passive voice could be seen in the control group, which is in line with previous research that learners significantly underused passives (Granger, 1997). In fact, there were three attempted uses of passives but all turned out to be unsuccessful due to the wrong choice of verbs (e.g., *An objection should not **be given** to welfare lottery any more*). The failure of using passives could be explained away by the fact that there were not adequate collocational patterns of *V + SN* stored in students' interlanguage system. When learners attempted to use the target word in writing, they tried to avoid employing

unfamiliar syntactic patterns and preferred using a ready-acquired structure. This is one of the possibilities that the two structures (i.e., *There BE + SN* and *Theme + Copular verb BE + SN*) were frequently used but few instances of *V+ SN* structure were observed in the control group's written texts. Clearly, inadequate knowledge of *V+SN* patterns deterred students from retrieving appropriate verbs to predicate the target nouns, which leads to more lexical collocation errors and fewer instances of *V + SN patterns*, particularly in passives.

A close analysis of the written texts produced by the two groups also shows that their tones as reflected in the use of SNs were vastly different. In an attempt to use *V + SN* patterns, the control group tended to begin the sentence with a collective noun *people*, pronouns *we* and *I*, which sounds less objective in tone and less formal in argumentative genre. However, in the experimental group, subjective interpersonal metaphors, such as *I* and *we*, were less used as *Subject*, thus reducing the involvement of subjectivity. This finding, to some degree, conforms to the findings by Herriman and Aronsson (2009), who, through comparing argumentative writing by Swedish advanced learners in the International Corpus of Learner English (*ICLE*) and native speakers in the Louvain Corpus of Native English Essays (*LOCNESS*), discovered that NNS students used more subjective interpersonal metaphors than NS students in English writing. So far, as the current study is concerned, the finding that more passive voices and less subjective interpersonal metaphors were used with the target nouns by the experimental group lends support to the argument that observing the recurrent language patterns from the corpus-informed concordances

enables learners to produce more native-like output not only in term of accuracy and complexity but also in the organization of information at discourse level.

On the whole, the above findings indicate that in L2 writing instruction, lexical collocations and syntactic structures are not two discrete but inter-related areas. The acquisition of collocations leads to improved use of syntactic patterns which facilitates learners to produce a variety of grammatical patterns. This evidence may provide useful information for language practitioners and L2 learners that lexis and grammar are two inseparable areas in language teaching and learning. Such a notion is in consonance with what has been advocated in linguistic research (e.g., Sinclair, 1991) that lexical studies must be incorporated with grammar research. Hunston and Francis (2000) also stated that a description of a word and its patterns “cannot be classified under the heading either of lexis or of grammar...[and] taking lexis as the starting point leads to a very different kind of grammar” (p.250). This in turn suggests that in language teaching and learning, particularly in L2 writing instruction, a word must be taught with not only its definition, spelling, parts of speech, but above all, the words it accompanies and grammatical structures which it is often fixed in. Clearly, materials generated from corpora in the form of concordances can provide effective scaffolding for learners to study lexis with a variety of its possible accompanying grammatical patterns in an orchestrated manner.

5.2.2. Encouraging usage-based learning

The findings of this study provide hard evidence that studying a word with in its KWIC form encourages learners to have a different perception of vocabulary learning

because it encourages usage-based learning. This echoes Johansson's (2009) suggestion that "the usage-based model and the relevance of corpora deserve to be recognized in works on second-language acquisition" (p.39). In second language acquisition, rule-based learning and usage-based learning are two dissimilar types of learning, which have been discussed along with other learning theories; i.e., deductive vs. inductive learning (Shaffer, 1989), and explicit vs. implicit learning (Levin, 1969; Scott, 1989). It has been acknowledged that the development of explicit rule-based knowledge is crucial in helping L2 learners sort out input and facilitate output (Ellis, 1993). Awareness of abstract rules or principles facilitates learners to produce language through analogy. Nevertheless, what should be pointed out here is that analogy is not an infallible guide to sentence building (Hornby, 1954, cited from Hunston & Francis, 2000). Mere emphasis on grammar rules in language teaching may lead learners to over-generalization of the acquired rules, which eventually results in language fossilization (Skehan, 1998).

In contrast to rule-based approach, usage-based approach was developed in light of the notion that L2 acquisition is basically instance-based or exemplar-based learning (Langacker, 2000). In Langacker's usage-based model, grammar is regarded as a network of exemplars and their schemas that emerge from actual usage, i.e., the network is gradually complexified through frequent experience of exemplars. This model emphasized the importance of semantics and blurred the boundary between lexicon and grammar. In line with Langacker's model, Hornby (1954) highlighted the importance of usage-based approach in SLA that "analysis [of rules] is helpful, but the learner is, or should be, more concerned with sentence-building. For this he needs to know the patterns of English

sentences and to be told which words enter into which patterns” (cited from Hunston & Francis, 2000, p.4).

Introducing corpus to language teaching allows learners to study the target language by adopting a usage-based approach. Corpus is a valuable resource giving practical guidance to language learners on usage rather than to expound analysis, and can be a concrete actualization for the rather vague “previous experience of language” (Hunston & Francis, 2000, p.3). Through observing the KWIC concordances, students may become aware of the typical and unusual syntactic patterning. This learning process helps students to become “rule & instance learners” (Ellis, 1993, see Abstract), who are able to see the rules applied to instances.

5.2.3. Raising learner awareness of collocation and colligation

According to the survey results, the concordance exercise was consistently emphasized by the students as the most beneficial resource for vocabulary learning, in particular collocation acquisition. In consonance with the findings of relevant previous research (Boulton, 2009), the qualitative analysis of lexico-grammatical features of the SNs in both groups’ written texts also suggests that corpus-informed concordances apparently had greater value in raising awareness of the co-occurring words associated with the target lexical items than traditional resources such as dictionaries and grammars.

After observing the given KWIC concordance lines, students were able to notice the lexico-grammatical patterns and produce appropriate collocations in the immediate posttest, which conformed to what is suggested by Hyland and Milton (1997) that KWIC

presentation can draw learners' attention to the target items, thus encouraging noticing.

Learning procedures must make patterns salient so as to help students notice patternings and to speculate about them (Wray, 2000). Such a type of learning process can be realized by the studying and analyzing corpus learning materials.

5.2.4. Retention of lexico-grammatical patterns

Examination of the occurrences of the target SNs in the delayed posttest (the free writing test) shows that the experimental group used far more SNs (51 correct occurrences) than the control group (9 correct occurrences), and was able to use the acquired patterns in their writing. This indicates that the concordance exercises, to some degree, facilitated learners to memorize the target patterns. The layout of KWIC concordance lines is an asset for learning lexico-grammatical patterns with the search word highlighted in the center and the co-text surrounded at its left and right. Such a list enables students to “look for eventual patterns in the surrounding co-text, which proffer clues to the use of the keyword item” (Partington, 1998, p.9).

From a psycholinguistic perspective, the role of patterns or prefabricated units is crucial for language acquisition and production. The process of language acquisition involves *pattern analysis*, *pattern imitation* and *pattern memorization* (Ellis, 1985). KWIC concordances offer an innovative way for learners to analyze the linguistic patterns, allowing them to get involved in a cognitive learning process as an active language analyst, whereas the traditional references such collocation dictionaries and grammars do not possess this merit as linguistic information is often presented in a straightforward manner.

In contrast to dictionary consultation, observing concordances encourages inductive learning (examples first, then rules) and discovery learning (Johns, 1991; Sun, 2000). This inductive learning process may lead to deeper cognitive processing; hence better retention of the target words than simply being taught their definition and usages (Cobb, 1997; Laufer & Hulstijn, 2001).

Although the experimental group greatly outperformed the control group in the retention of the target words, the repeated errors in the delayed posttest by certain students in the experimental group are noteworthy. Several students made an error in the pretest and corrected it in the immediate posttest. After a two-week interval, the same error in the pretest re-occurred in the delayed posttest. In view of this, we cannot claim that there is an absolute correlation between corpus use and retention of target lexico-grammatical patterns. More controlled experiments in relation to corpus use and retention of linguistic patterns are necessary in order to provide explicit evidence to support the claim that corpus is helpful for retention of language patterns. As a closing remark to Section 5.2, a diagram is presented below (see Figure 5.1) to symbolize the effects of corpus-informed approach to vocabulary use in L2 writing.

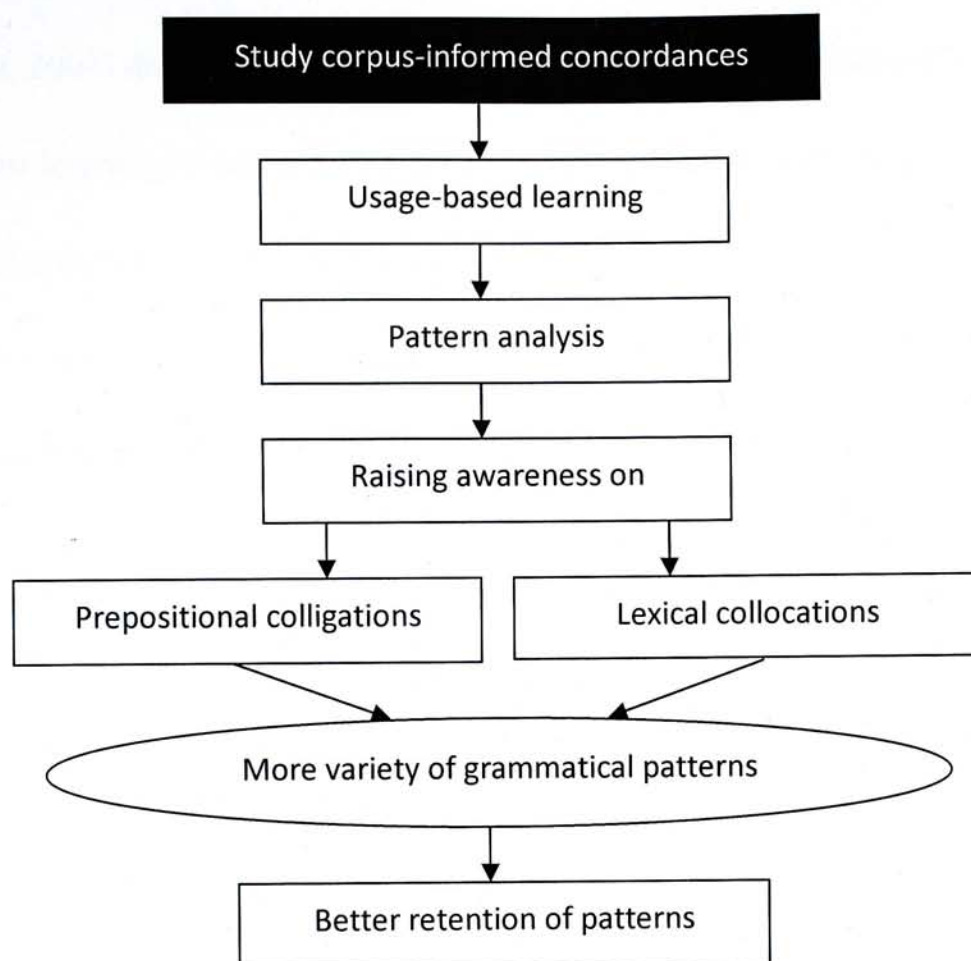


Figure 5.1. Effects of corpus-informed approach on vocabulary use in L2 writing

5.3. Did students think that corpus-informed approach helped their writing?

Generally, most students had a positive response to corpus use on their writing. It is worth noting that all the informants concurred that concordances were most useful in vocabulary learning, especially in improving collocational use. However, in relation to content schemata input, students' responses appeared to be less positive due to the truncated sentences and the cultural-load information.

In regards to students' evaluations of concordances on grammar patterns acquisition, it is interesting to find that students had a favorable response (mean score of 5.90), which ranks No. 2 following vocabulary learning (mean score of 6.30). This result is different from what was identified in previous research (Vannestal & Lindquist, 2007; Yoon &

Hirvela, 2004) that students did not rate corpus as a favorable and useful resource for grammar learning. Yoon and Hirvela attributed this unfavorable response to the students' unfamiliarity with the notion of lexico-grammatical patterning: "grammar as presented through a corpus orientation did not conform to the notion of learning about grammar" (Yoon & Hirvela, 2004, p.269). However, in this study, students' favourable response in relation to the usefulness of concordances for grammar acquisition appears to challenge this claim. To a certain extent, this suggests that observing a key word in concordances helps students become aware of the close relation between lexis and grammar patterns. If this is true, concordances can be valuable resources for helping students to become aware of the notion that "lexical patterning is seen as the key to grammatical description" (Owen, 1993, p.169).

Also worth noting is that students' responses to the retention of target language patterns are varied greatly. Quite a number of students took a conservative approach and did not perceive that they were able to memorize the acquired patterns and use them in future writing. This is a subtle reminder that although corpus activity may facilitate learners to memorize linguistic patterns by offering them a chance of inductive *pattern analysis*, corpus activity alone cannot automatically help students to have a better retention of the sentence patterns. Follow-up learning activities, such as guided writing tasks, can be conducted in order to offer learners a chance for *pattern imitation*, and eventually lead to *pattern memorization* and *pattern internalization*.

Students' evaluation regarding the usefulness of corpus for word retention contradicted with the textual results in their written texts, revealing that a great number of

students were not confident in using these words and the acquired patterns in future writing although they could use them in a short interval (2 weeks). However, whether or not these students could use these words and their patterns in a longer run is still unknown.

In short, the findings that more occurrences of target words were observed in the experimental group suggest that concordance exercises facilitated students to have a better retention of the target words and their patterns as compared with other forms of consultation materials, such as dictionaries. What should be pointed out, though, is that this controlled experiment was conducted within a relatively short period of time; thus more research is needed to substantiate the claim if it is to be generated to other contexts where students are of a different nature or experiments have a longer time span.

5.4. Towards a tentative model of corpus-informed writing instruction

The role of small corpora in L2 writing has been discussed in English language teaching along with writing materials development (e.g., Coxhead & Byrd, 2007; L. Flowerdew, 2010; Tribble, 2001). Tribble (2001) has provided a useful framework on how a genre-specific micro corpus can be incorporated into L2 writing classroom to improve learners writing performance. According to Tribble (2001):

Writers need at least four kinds of knowledge when they approach a writing task: knowledge of content, writing process, context, and language system. An appropriate corpus of exemplars of the genre in which learners are interested provides them with an important resource, as it offers essential information about the lexico-grammar, patterning and organization which have made these texts allowable contributions to the genre. (p.382)

Key words generated from a small corpus which deals with same content will help learners to gain content knowledge, i.e., knowledge of the concepts involved in the subject area (Tribble, 1997). Analyzing concordances assists learners to understand linguistic behaviors of a lexical item in multiple contexts. Such lexical input can provide significant assistance to learners when attempting to produce a written text (Tribble, 2009).

Through incorporating the findings of the present study with the framework of genre-specific corpus approach to L2 writing established by Tribble (2001), previous accounts of concordance learning (Sun, 2003) and learning materials developing for L2 writing (Coxhead, 2007), a tentative model of corpus-informed writing instruction is proposed with the attempt to describe: 1) the role of teachers and students in this type of writing instruction; 2) the procedures of preparing learning materials; and 3) the implementation in classroom. Previous studies on data-driven learning emphasized the role of learners as linguistic researcher or “Sherlock Holmes” (Johns, 2000, p.108). Within this theoretical framework, the role of teachers appeared to be underestimated. Later on, empirical research found that teachers’ intervention is crucial in facilitating learners to successfully formulate linguistic patterns or achieve learning goals (Carter & McCarthy, 1995; L. Flowerdew, 2009; Sun, 2000). As a matter of fact, teachers play a vital role in ensuring learners to achieve learning objectives through data-driven learning. Teachers “have the responsibility to create opportunities in which learners can develop and extend the process, context and language system knowledge they need in order to write successful texts” (Tribble, 2001, p.382). They play multiple roles at different stages as a writing task

organizer, a linguistic researcher and an activity facilitator. Without teachers' intervention and organization, exposure to a large amount of cut-off sentences is likely to be a fear for students that will frustrate them in language learning.

5.4.1. Preparing Materials

5.4.1.1. Establishing a topic-based corpus

It is equally important to have teacher intervention before data-driven learning activities. Teachers need to choose the appropriate corpus, to sift the concordance output, establish topic-specific corpora by collecting appropriate texts from the Internet without violating copyrights, download texts and convert them into plain texts.

5.4.1.2. Selecting target lexical items

Use corpus-analysis tool to generate a word list. Select target lexical items according to word frequency. "A frequency sorted wordlist offers an effective way of beginning to study a text or collection of texts, as it gives an immediate insight into words which appear with high frequency" (Tribble, 1997). However, it should be noted here that frequency is not the only criterion for words selection as most of the top high frequency words are grammatical words. At this point, lower frequency words also need to be taken into account and observed. The selected words for studying could be content schemata and problematic words in relation to collocation and colligation use. To do this, NNS teachers' intuition and teaching experience are of vital importance.

5.4.1.3. Analyzing textual patterns

Based on the comments of the independent raters, one of the problems in the students' texts is lack of coherence and cohesion. It is worthwhile for teachers to observe each single text to analyze the textual patterns or structural organization in order to discover the typical linguistic patterns which might help learners organize text structures (Tribble, 2001). Alternatively, clusters can be generated by corpus analysis tools. Analysis of organizational patterns can help learners present their ideas logically.

5.4.1.4. Designing a web-based learning platform

Design a web-based interface for students to explore the small topic-based corpus under teachers' instructions. Ensure learners can go further to larger contexts when they read the truncated lines of the target words. The outcome of a reading activity can serve as input for writing (Grabe & Kaplan, 1996). Through exploring the corpus, students can have a chance to carry out extensive and intensive reading activities as they are exposed to substantial amount of data with texts of the same topic. Reading on the same subject helps learners obtain enough linguistic knowledge necessary for their writing.

5.4.2. During the exploration of a topic-specific corpus

Previous studies on using corpus in language classroom (e.g., Thurstun & Candlin, 1998; Vannestal & Lindquist, 2007) found that exploring a corpus or data-driven learning was frequently reported as a boring process. This current research also found that most of the students hold a reserved attitude towards data-driven learning, commenting that it was quite

boring to study merely a few words with so many cut-off sentences. Appropriate guidance needs to be provided to students when they come across difficulties or get confused in analyzing the concordances. One of the students in this present study commented on her confusion when studying the concordances:

We don't know the difference such as by using different prepositions. Some words can collocate with different prepositions. The concordance lines do not tell us in what circumstance certain preposition should be used.

Thus, as L. Flowerdew (2009) proposed, corpus can be used as learning resources for checking vocabulary use in the classroom, but teachers' intervention is necessary in data-driven learning.

5.4.3. Follow-up activities after exploration of the topic-specific corpus

Peer or teacher-students discussions on what learners discover after the observation of the concordances need to be carried out in classroom after exploring the corpus. It is essential that students have an opportunity to discuss on how to organize their ideas with content schemata logically. Guidance from teachers will definitely play a positive role in helping learners to link the ideas or content schemata with coherence and logic. Otherwise, the content schemata input will turn out to have a negative effect on content organization as is evidenced in the current research.

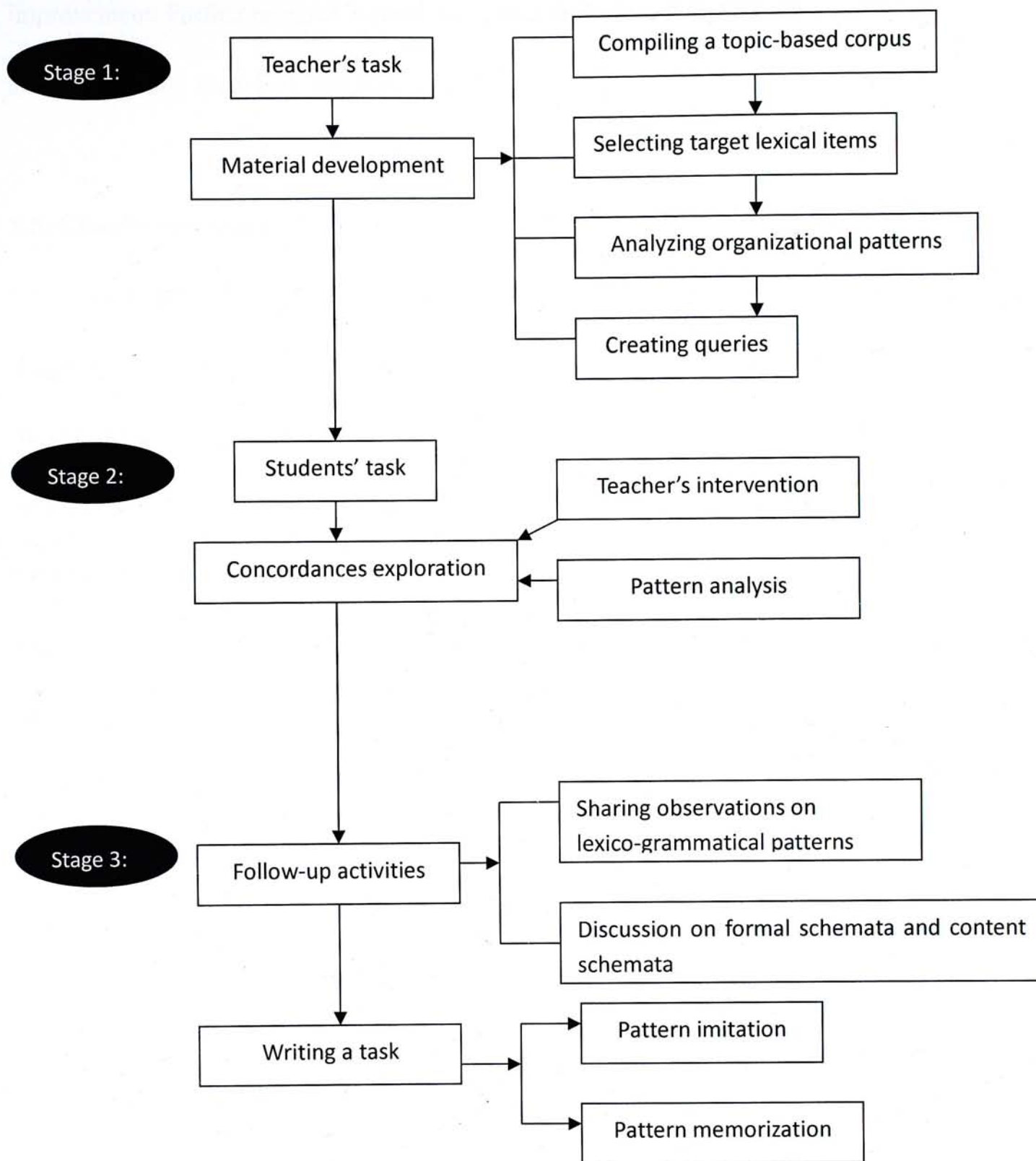


Figure 5.2. A tentative model of corpus-informed approach to L2 writing

The current tentative model is developed by adapting and integrating Tribble's (2001) framework with the findings of the current study. It is dynamic and open for further

improvement. Further research is needed in order to obtain solid evidence on its feasibility and applicability in an EFL and ESL context.

5.5. Chapter summary

This chapter has discussed the effects of corpus-informed concordance on L2 writing. To be more specific, it examined whether and the extent to which corpus-informed lexical input (SNs and CSNs) had impact on students' overall writing performance, realization of collocational patterns, syntactical complexity and retention of the acquired linguistic patterns. Based on students' written output, it can be seen that corpus plays a significant role in helping learners in the areas of lexical collocations and prepositional colligations. Most importantly, improvement of collocational use greatly facilitates learners to acquire a variety of grammatical structures. On the whole, in light of the findings obtained from this study, it is arguable that corpus-informed concordances are an efficacious resource for improving vocabulary use in students' written texts.

CHAPTER SIX CONCLUSION

This chapter begins with a brief summary of the current research by restating its significant findings; followed by a discussion on its implications for L2 writing pedagogy. At the end of the chapter, limitations and suggestions for future research on the application of corpora to L2 writing instruction are also presented.

6.1. Summary of this study

Many researchers agree that linguistic discrepancy, manifested as contrastive rhetoric, is a common difficulty for L2 learners in their writing (e.g. Engber, 1995; Hinkel, 2004; Leki & Carson, 1994) and that a lack of vocabulary knowledge can prevent students from making progress on L2 writing (Cobb & Horst, 2001). A corpus-informed approach to helping learners' L2 writing has been advocated in recent years (Lee & Swales, 2006; McCarthy, 2008; Tribble, 1997, 2002). On the basis of a detailed literature review presented earlier in Chapter Two, this empirical study attempted to investigate whether and in what ways corpus-informed learning materials have an influence on learner's argumentative writing. Through adopting both qualitative and quantitative analysis of students' written texts and students' evaluations of corpus use, four significant findings were identified. They are presented as follows.

6.1.1. *Enhancement of lexico-grammatical patterns*

Use of corpus-informed materials can enhance the acquisition and production of lexico-grammatical patterns. As reported in Chapter 4, in the immediate posttest, written

output of the experimental group contained a higher variety of grammatical structures and fewer errors in using the target lexical items after the observation of concordances, whereas the control group did not make noticeable progress in using the target words. This may suggest that students' conventional view of vocabulary learning in which a word is learned only with its definition, spelling and part of speech seems inadequate and detached from context, particularly when the words are used as productive vocabulary in L2 writing. The reason why the control group did not make much progress in using the target SNs even after they were directed to consult dictionaries can partially be attributed to fact that dictionary consultation may not consciously arouse their attention to notice which words go with which. Thus, when drafting an essay, they may have been confronted with a problem described by Lewis (2000):

[Students] may know quite a lot of individual words which they struggle to use, along with their grammatical knowledge, but they lack the ability to use these words in a range of collocations which pack more meaning into what they say. (p.14)

In many cases, the sentences produced by the control group were grammatically accurate, but the whole sentence sounds inadequate in terms of fluency and coherence due to the choice of an inappropriate verb to predicate the target SN. This written output which assumes an insufficient interlanguage nature indicated that the "*slot and filler* approach to the teaching grammar and vocabulary has not sensitized students to the collocational constraints on word combinations" (Wooland, 2000, p.30).

The corpus-informed approach, on the other hand, integrates both lexical and grammatical collocations by presenting the sentence patterns where a target word is

actually used in context. Concordances highlight language patterns and help L2 learners notice the lexico-grammatical patterns of the target words, which is not often achieved by the conventional teaching approach of focusing on structural rules. The advantage of corpus-studies of an individual word is that it includes both “lexical and grammatical associations” (Biber, 1999, p.290), where “the lexical associations are the tendencies for the target grammatical construction to co-occur with particular words.”

6.1.2. Enhanced awareness of the importance of collocations

Corpus-informed exercises enhance learners’ awareness of the importance of collocations in L2 writing. KWIC concordances not only offer students effective scaffolding for analyzing patterns, but also play a significant role in helping them develop independent learning strategies including noticing and discovery learning. Through constant exposure to the regular patterning of word usage provided by corpus-informed materials, it is believed that students will also raise their consciousness in their daily learning of vocabulary. In other words, the corpus-informed approach can benefit and contribute to non-corpus-based regular learning. The traditional meaning-based learning can be compensated, to a certain extent, by the enhanced conscious noticing, learning, and eventually internalization of language forms. In this sense, the use of KWIC concordances can help develop and raise learner autonomy in learning productive vocabulary.

6.1.3. Pivotal role of prior grammatical knowledge in corpus-informed learning

Prior knowledge of grammatical rules plays a pivotal role in achieving the effectiveness of concordance-based observation. The findings proved that concordance

observation encourages usage-based learning, which allows learners to learn language by analyzing fabricated units rather than breaking down the linguistic patterns into individual words. However, when we emphasize the benefits of usage-based learning, we cannot underestimate the role of rule-based knowledge in corpus activities. Usage-based learning and rule-based learning are, virtually, not two non-compatible dichotomies; in fact, they are reciprocal approaches required in the process of interpreting word usages from concordance data. Rule-based system can facilitate students to analyze patterns and word usage successfully. Lack of grammatical rules knowledge may lead to the failure of formulating linguistic rules and patterns from the concordances. This can be one of the causes why a small number of students in the experimental group were unable to use the target word accurately in their writing after doing the concordance activities. Students with inadequate knowledge of abstract rules would often find it hard to recognize the boundary of chunks, i.e., which word is the “real friend” accompanying the observed word. It is possible that students mistakenly judged the word clusters by simply chunking the target word with its first immediate word either to the right or the left, when in fact these words are “false neighbors” which belong to other constituents of the sentence. Thus, simply directing students to observe the target words on the left and right sides sometimes may be misleading, especially in cases where learners are not equipped with proper grammatical knowledge to correctly judge and formulate the patterns.

6.1.4. Insignificant correlation between learning CSNs and ideas development

The keyword function of corpus analysis tools allows us to generate a list of content schemata words related to the prescribed topic from a topic-specific corpus; however, students' exposure to concordances of those content schemata words does not automatically help them enhance their writing content due to several reasons. First, the cut-off concordance lines deter students from obtaining information embedded with the target content schemata words. Second, the culture-loaded concordance lines exacerbate the comprehension of the incomplete sentences. In light of this, full contextual use of the words needs to be provided to students for further comprehension. Third, merely providing students with key words related to the topic is not adequate for content development if students have not sufficient knowledge and awareness of using coherent and cohesive devices to present their ideas critically and logically.

6.2. Pedagogical implications

The findings of this study provided some useful implications for teaching L2 writing, especially where corpora are integrated into writing instruction in the EFL context. The significant progress in using the target words in students' writing along with their overall positive attitudes towards corpus use testified to the feasibility and usefulness of corpus-informed approach to L2 writing instruction. Its benefits to L2 writing pedagogy can be twofold.

6.2.1. Writing materials development

This study provides some practical guidelines to writing teachers, especially those NNS teachers in an EFL context, on how to develop corpus-informed writing materials to enrich learning resources. The procedures of developing corpus-informed materials presented earlier in Chapter 3 can be employed in the actual writing classroom. There are four steps to do this: 1) establishing a topic-specific corpus; 2) selecting target words, 3) sifting concordance lines and 4) formulating queries.

For establishing a topic-specific corpus, Nelson's (2009) approach of collecting online texts provides a technically viable method for obtaining quality texts from the Internet in an effective and economical manner without violating copyrights. In selecting the target words, two major types of words can be taken into consideration. The first type is those words that learners often have problem in using their proper collocation and colligation. The second type is content schemata words related to the prescribed topic. The reason for selecting the problematic words is to encourage a need-driven learning. Only when students notice the gap between the target language and their interlanguage will they be motivated to observe the lines and do the corpus-informed tasks. In addition, vocabulary learning is not just about learning new words, but also learning the partially acquired words, especially the knowledge of their lexico-grammatical patterns. Concordances can be efficacious resources for enriching word usage knowledge.

As regards to sifting concordances, the lines containing recurrent patterns can be incorporated into learning materials. To ensure that the common linguistic patterns are presented in the exercises, the use of traditional references, such as collocation dictionaries,

is necessary in the sifting process. The role of traditional resources should not be underestimated when using corpus for language teaching purpose. As O’Keeffe, McCarthy and Carter (2007) have stated:

The advantage of looking up a lexico-grammatical query in a corpus is that it provides us with many examples of the search item in its context of use. However, a corpus will not tell us the meaning of word or phrase. This is something that we have to deduce from the many examples that are generated. Combining a dictionary and a corpus can be a valuable route in a pedagogical context. (p.3)

6.2.2. Implementation of corpus-informed activities

For studying the lexico-grammatical patterns of problematic words, paper-based print-outs with cut-off sentences can be quite efficient for raising learners’ awareness of word combinations. However, for content schemata words, a fuller context of the target word can be presented to offset the negative influences of cut-off sentences which seemed to hamper the development of ideas in writing. Therefore, full contexts of target words must be made available to learners for better understanding the contextual use of the words related to the writing subject. If it is possible, online concordances with accessibility of full context rather than mere presentation of truncated lines can be implemented as they seem to be a more desirable approach for conducting corpus-informed activities, in particular the studying of content schemata words.

The corpus exploration process allows learners to obtain vocabulary and patterns related to the writing topic through intensive and extensive reading. Concordance-based studies of content schemata words can be a short-cut for learners to read a sufficient

number of texts relevant to the intended writing topic. Directing learners to observe the KWIC concordances of content schemata words can serve as a starting point of corpus-informed reading activity. Through reading a fuller context, students are given an opportunity to efficiently read related texts relevant to the writing topic and obtain more relevant ideas, which otherwise would cost them more time using traditional instruction materials such as textbooks or handouts of writing samples.

However, it is worth pointing out that learners' proficiency level may affect the effectiveness of content schemata input. It is likely that this type of input may be more appreciated by low level students since learners at this level usually have a smaller size of vocabulary, and tend to struggle with retrieving words related to the writing topics. The input of content schemata words will help them reduce time for choosing the desired words and enable them to better express their ideas. On the other hand, students with relatively higher language proficiency, such as upper-intermediate-level students targeted in this study, may be less motivated to study the target content schemata words since they have a comparatively bigger size of vocabulary to express their ideas. For these students, providing them with basic vocabulary related to the prescribed writing topic may not be very helpful for their L2 writing. Rather, they may prefer learning some cohesive or rhetorical devices in order to polish their writing contents.

To sum up, incorporation of corpus into writing instruction needs to be pedagogically mediated. In corpus-informed writing instruction, teachers play a central role from the initial stage of material development to the implementation of corpus-related learning activities. For practitioners, the statement by Johansson (2009) can be used as a gentle

reminder that “corpora are important in basic research, and they have a role to play in the classroom as well. But let’s not exaggerate. Corpora are no replacement for natural communication. They cannot replace the teacher.” (p.42)

6.3. Limitations and suggestions

This experimental study attempted to investigate the effects of corpus-informed approach to L2 writing in terms of writing quality and language use through examining the written output of two groups of students. Factors related to L2 writing, such as writing skills and writing strategies, were not within the research scope of this study. In terms of research design and methodology, the present study is limited in several ways. The limitations are illustrated below; along with some suggestions for possible replications of this type of studies in the future.

6.3.1. A longer experimental time frame

An inherent limitation of this controlled experiment is that it only investigated the short-term effects of corpus-informed approach to L2 writing within a fixed period of time. As we have seen, four weeks was not likely to be sufficient to detect learners’ writing development in terms of language use. Certainly longitudinal designs have more merits in investigating these issues regarding progress making in L2 writing. To a certain extent, the study only indicated that the students were able to use the acquired patterns in their writing right after the treatment and two weeks later in the delayed posttest. However, whether they can use these newly acquired patterns in future writing after a longer time interval, e.g., at

the end of a school term or two months after the treatment, remains unknown. It is therefore quite worthwhile to extend the experimental time frame in order to examine its long-term effect.

6.3.2. *More lexical input*

If we refer back to the research design described earlier in Chapter 3, we can see that this research adopted a pretest-posttest design with only one treatment, i.e., the participant receiving lexical input once. This could possibly explain the reasons underlying the insignificant change in terms of overall writing quality because students were provided with an input of only 10 words. Clearly, if a longitudinal study had been possible, more writing tasks could have been designed and conducted so that students could have received more lexical input through exploring corpus-informed concordances before each writing task. With more lexical input, the effects of corpus use on overall writing quality would probably be different from what is found by the present study.

6.3.3. *More comparison groups*

This study used two groups of students, the experimental group receiving concordance input while the control group used dictionary consultation. It would be interesting to investigate and compare the results of L2 writing with more groups in different modes of learning, such as learning through paper-based concordances, online-concordances, dictionary consultation, corpus plus dictionaries or corpus only, with teachers' intervention or without teachers' intervention.

6.3.4. Different proficiency levels

This study only focused on a group of learners with one proficiency level (upper-intermediate). It will be useful to conduct further research, aiming at learners with different proficiency levels (from pre-intermediate to advanced levels), to investigate their perceptions of corpus use for L2 writing, and linguistic progress in their written texts after studying corpus-informed learning materials.

6.3.5. Web-based concordances and more follow-up learning activities

This study adopted paper-based design of concordance learning materials rather than direct application of electronic corpora. As listed in Boulton (2008), using print-outs has a few advantages: no need for a computer laboratory and no technical problems and training required during hands-on practice. For collocation studies, print-outs with truncated concordances lines appear to be quite effective. However, for L2 writing, cut-off sentences seem to be the major problem affecting their writing content development. It is more advisable that learners read the full contextual use of the target words in order to better understand the information and gain more ideas of the writing subject. Thus, for content schemata input, online-concordances should ideally be provided to learners in order to ensure that they can get access to a fuller context. It would be worthwhile to conduct further research on comparing the effects of online concordance exploration of content schemata on L2 writing quality with the one using mere observation of paper-based exercises with cut-off lines.

6.3.6. Case studies

This study adopted a controlled experiment with a relatively medium size of population to investigate the short-term effect of corpus-informed approach to L2 writing. However, this empirical study has provided some information of and insights into corpus-informed writing instruction, which can probably be subsequently applied to the similar studies aiming at investigating the longer-term effects of corpus application, the use of a qualitative research method, such as case studies are equally desirable as they will allow us, as Braine (2002) has stressed, “[to] get richer information about the learners, their learning strategies, how their attitudes, personalities and goals interact with the learning environment and the nature of their linguistic growth” (p.66). By doing case studies, more naturally occurring data of and in-depth insights into the development of learner writing literacy can be obtained to complement the findings of controlled experimental research.

6.4. Closing remarks

To conclude, what J. Flowerdew (1996) noted more than a decade ago is still of value for research to date on corpus and language teaching:

There is a danger of the enthusiasm for concordancing being inflated to such an extent that concordancing is seen as a sort of language teaching panacea. Carefully conducted evaluative studies will ensure that such an inflated view will not prevail. Instead of concordancing being promoted as a panacea, considered evaluative studies will allow concordancing to be incorporated appropriately into the teacher’s battery of reference and teaching resources as the useful additional teaching and learning tool that it undoubtedly is. (p.112)

Bibliography

- Aston, G. (2004). Corpus upon corpus: a bout of indigestion?. Paper presented at the Sixth Teaching and Language Corpora Conference, Granada Spain, 2-9 July.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). New York: The Edwin Mellen Press.
- Biber, G. (1999). Investigating language use through corpus-based analyses of association patterns. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 287-315). Stanford, CA: CSLI Publications.
- Boulton, A. (2008). DDL: Reaching the parts other teaching can't reach?. In A. Frankenburg-Garcia (Ed.), *Proceedings of the 8th Teaching and Language Corpora Conference* (pp. 38-44). Lisbon, Portugal: Associação de Estudos e de Investigação Científica do ISLA-Lisboa.
- Boulton, A. (2009). Testing the limits of data-driven learning: language proficiency and training. *ReCALL*, 21(1), 37-54.
- Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60 (3), 1-39.
- Braine, G. (2001). A study of English as a foreign language (EFL) writers on a local area network (LAN) and in traditional classes. *Computers and Composition*, 18, 275-292.
- Braine, G. (2002). Academic literacy and the nonnative speaker graduate student. *Journal of English for Academic Purposes*, 1, 59-68.

- Braun, S. (2007). Designing and exploiting small multimedia corpora for autonomous learning and teaching. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom: Selected papers from TaLC 2004* (pp.31-46). Amsterdam: Rodopi.
- Carter, R. & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, 16 (2), 141-158.
- Chambers, A. & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL*, 16(1), 158-172.
- Cobb, T. (1997). Is there any measurable learning from hands on concordancing?. *System*, 25(3), 301-315.
- Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), 345 - 360.
- Cobb, T. & Horst, M. (2001). Reading academic English: carrying learners across the lexical threshold. In J. Flowerdew & A. Peacock (Eds.), *The English for academic purposes curriculum* (pp.315–329). Cambridge: Cambridge University Press.
- Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52(1), 57-63.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Coxhead, A. & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16, 129-147.
- Coxhead, A. (2008). Phraseology and English for academic purposes. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language teaching* (pp.149-161). Amsterdam/Philadelphia: John Benjamins.

- Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, 27, 91-113.
- Engber, C.A. (1995). The relationships of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43.
- Flower, L. & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Flowerdew, L. (2002). Corpus-based analyses in EAP. In J. Flowerdew (Ed.), *Academic discourse* (pp. 95-114). Harlow, UK: Longman.
- Flowerdew, L. (2008). *Corpus-based analyses of the problem-solution pattern: a phraseological approach*. Amsterdam/Philadelphia: John Benjamins.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393-417.
- Flowerdew, L. (2010). Using corpora for writing instruction. In M. McCarthy & A. O’Keeffe (Eds.), *The Routledge handbook of corpus linguistics* (pp. 444-457). London: Routledge.
- Flowerdew, J. (1996). Concordancing in language learning. In M. Pennington (Ed.), *The power of CALL* (pp. 97-113). Houston, TX: Athelstan.

- Flowerdew, J. (2003). Signalling nouns in discourse. *English for Specific Purposes*, 22, 329-346.
- Flowerdew, J. (2009). Use of signalling nouns in a learner corpus. In J. Flowerdew & M. Malhberg (Eds.), *Lexical cohesion and corpus linguistics* (pp.85-102). Amsterdam/Philadelphia: John Benjamins.
- Flowerdew, J. (2010). Use of signalling nouns across L1 and L2 writer corpora. *International Journal of Corpus Linguistics*, 15(1), 36–55.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. London: Longman.
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells?. *TESL-EJ* 8(4), A1, 1-37. Retrieved from <http://www.tesl-ej.org/wordpress/>.
- Gaskell, D. & Cobb, T. (2004). Can learners use concordance feedback for writing errors. *System*, 32(3), 301-319.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam/Philadelphia: John Benjamins.
- Granger, S. (1997). Automated Retrieval of Passives from Native and Learner Corpora: Precision and Recall. *Journal of English Linguistics*, 25, 365-374.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer (Ed.), *Corpora and language teaching* (pp.13-32). Amsterdam/Philadelphia: John Benjamins.
- Hanson-Smith, E. (1993). Dancing with concordances. *CAELL Journal* 4(2), 40.

- Herriman, J. & Aronsson, M.B. (2009). Themes in Swedish advanced learners' writing in English. In K. Aijmer (Ed.), *Corpora and language teaching* (pp.101-120). Amsterdam/Philadelphia: John Benjamins.
- Hinkel, E. (2004). *Teaching academic ESL writing: practical techniques in vocabulary and grammar*. London: Routledge.
- Hornby, A. S. (1954). *A guide to patterns and usage in English*. Oxford: Oxford University Press.
- Hymes, D. (1973). On communicative competence. In J.B. Pride & J. Homes (Eds.), *Sociolinguistics*. Harmondsworth: Penguin.
- Hyland, K. & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Second Language Writing*, 6(2), 183-205.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- Hunston, S. & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: Benjamins.
- Johansson, S. (2009). Some thoughts on corpora and second language acquisition. In K. Aijmer (Ed.), *Corpora and language teaching* (pp.33-45). Amsterdam/Philadelphia: John Benjamins.
- Johns, T. (1986). Micro-Concord: A language learner's research tool. *System*, 14 (2), 151-162.
- Johns, T. (1990). From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning. *CALL Austria*, 10, 14-34.

- Johns, T. & King, P. (Eds) (1991). *Classroom concordancing: Special Issue of ELR Journal 4*. University of Birmingham: Centre for English Language Studies.
- Johns, T. (2000). Data-driven learning: the perpetual challenge. In K. Bernhard & G. Marko (Eds.), *Language and computers, teaching and learning by doing corpus analysis: Proceedings of the fourth international conference on teaching and language corpora* (pp. 107-117). Graz 19-24 July: Rodopi.
- Johnson, C. (2001). Book reviews, Pattern grammar: A corpus-driven approach to the lexical grammar of English. *Computational Linguistics*, 27(2), 318-320.
- Koo, K. (2006). Effects of using corpora and online reference tools on foreign language writing: A study of Korean learners of English as a second language, unpublished PhD thesis, The University of Iowa.
- Krishnamurthy, R. & Kosema, I. (2007). Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes*, 6(4), 356-374.
- Langacker, R. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1-63). Stanford, CA: CSLI Publications.
- Laufer, B. & Hulstijn, J.H. (2001). Some empirical evidence in the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 538-558.
- Laufer, B. & Nation, P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Leki, I. & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28, 81-101.

- Lee, D. & Swales, J.M. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self compiled corpora. *English for Specific Purposes*, 25(1), 56-75.
- Leech, G. & Candlin, C.N. (1986). *Computers in English language teaching and research*. London: Longman.
- Levin, L. (1969). *Implicit and explicit: A synopsis of three parallel experiments in applied psycholinguistics. Assessing different methods of teaching grammatical structures in English as a foreign language* (Research Bulletin No. 1). Gothenburg, Sweden: University of Gothenburg, Gothenburg School of Education, Department of Educational Research.
- Levy, M. (1990). Concordances and their integration into a word-processing environment for language learners. *System*, 18(2), 177-188.
- Lewis, M. (2000). There is nothing as practical as a good theory. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp.10-27). Hove, England: Language Teaching Publications.
- McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, 41(4), 563-574.
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung & S. Petch (Eds), *Computer learner corpora, second language acquisition and foreign language teaching* (pp.119-142). Amsterdam: John Benjamins.

- Milton, J. (2006). Resource-rich web-based feedback: Helping learners become independent writers. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 123-139). Cambridge: Cambridge University Press.
- Nation, P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nelson, G. (2005). Description and prescription. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp.460-465). Oxford: Elsevier.
- Nelson, G. (2009). Comparing Englishes: ICE and ICELite. *Proceedings of the 6th national natural language processing research symposium*. De La Salle University, Manila, 12-17.
- Noguchi, J. (2004). A genre analysis and mini-corpora approach to support professional writing by non-native speakers. *English Corpus Studies*, 11, 101-110.
- O’Keeffe, A., McCarthy, M. & Carter, R. (2007). *From corpus to classroom: language use and language teaching*. New York: Cambridge University Press.
- O’Sullivan, Í. & Chambers, A. (2006). Learners’ writing skills in French: corpus consultation and learner evaluation. *Journal of Second Language Writing*, 15(1), 49-68.
- O’Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, 19 (3), 249-261.
- Owen, C. (1993). Corpus-based grammar and the Heineken effect: Lexico-grammatical description for language learners. *Applied Linguistics*, 14(2), 167-187.

- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. Amsterdam: John Benjamins.
- Rayson, P. (2002). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*, PhD. Thesis, Lancaster University.
- Read, J. (2004). Research in teaching vocabulary. *Annual review of Applied Linguistics*, 24, 146-161.
- Schaffer, M.E. (1989). Are profit maximizers the best survivors?. *Journal of Economic Behavior and Organization*, 12, 29-45.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Scott, V. (1989). An empirical study of explicit and implicit teaching strategies in French. *The Modern Language Journal*, 73, 14-22.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stevens, V. (1991). Concordance-based vocabulary exercises: A viable alternative gap fillers. In T. Johns & P. King (Eds.), *Classroom concordancing: English language research journal 4* (pp.47-63). Birmingham, England: Centre for English Language Studies, University of Birmingham.
- Sun, Y-C. (2000). Using on-line corpus to facilitate language learning. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages. British Columbia, Canada.

- Sun, Y-C. (2003). Learning process, strategies and Web-based concordancers: A case-study. *British Journal of Educational Technology*, 34(5), 601-613.
- Thompson, P. (2001). Review of exploring academic English. *Language Learning & Teaching*, 5 (3), 28-31.
- Tribble, C. (1997). Improvising corpora for ELT: quick-and-dirty ways of developing corpora for language teaching. In J. Melia & B. Lewandowska-Tomaszczyk (Eds.), *PALC '97 proceedings* (pp.106-117). Lodz: Lodz University Press.
- Tribble, C. (2001). Small corpora and teaching writing: Towards a corpus-informed pedagogy of writing. In M. Ghadessy, A. Henry & R. L. Roseberry (Eds.), *Small corpus studies and ELT: Theory and practice* (pp.381-408). Amsterdam/Philadelphia: John Benjamins.
- Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 131-149). Harlow, UK: Longman.
- Tribble, C. (2009). Writing academic English: A survey review of current published resources. *ELT Journal*, 63 (4), 400-417.
- Tribble, C. & Johnes, G. (1997). *Concordances in the classroom: a resource guide for teachers*. Houston, TX: Athelstan Publications.
- Thurstun, J. (1996). Teaching the vocabulary of academic English via concordances. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, Chicago.
- Thurstun, J. & Candlin, C. (1997). *Exploring academic English: A workbook for student essay writing*. Sydney: NCELTR.

- Thurstun, J. & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267-280.
- Vannestal, E. & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 19(3), 329-350.
- Weber, J.J. (2001). A concordance-and genre-informed approach to ESP essay writing. *ELT Journal*, 55(1), 14-20.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Widdowson, H.G. (2002). Corpora and language teaching tomorrow. Keynote lecture delivered at the Fifth Teaching and Language Corpora Conference. Bertinoro, Italy, 29 July.
- Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics*, 21(4), 463-489.
- Wooland, G. (2000). Collocation: Encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp.28-46). Hove, England: Language Teaching Publications.
- Yoon, H. & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.
- Yoon, H. (2005). An investigation of students' experience with corpus technology in second language academic writing, unpublished PhD thesis, the Ohio State University.

Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology*, 12(2), 31-48.

Appendix One Questionnaire One

Part One: General information on English writing

Background information

1. Name in Chinese: _____
2. Class: _____
3. Student ID No.: _____
4. Date of Birth: _____
5. Gender: male _____ female _____
6. Email address: _____ (most applicable)
7. QQ: _____
8. Mobile Phone Number: _____ Flat phone Number: _____

General information on English writing learning

9. How many years have you been studying English _____
10. How do you rate your **writing** proficiency in English as compared with the proficiency of *the other students* in your class? (Circle one)
1. Excellent 2. Good 3. Fair 4. Poor
11. How important is it for you to become proficient in English writing? (Circle one)
1. Very important 2. Important 3. Not so important 4. Not important
12. Do you enjoy English writing? (Circle one)
1. Very much 2. Fair 3. A little 4. Very little
13. Why do you want to learn English writing? (Circle those that apply)
1. Interested in writing
2. Need it to pass the exams
3. Need it for my future career
4. Others (please list)

14. What types of writing do you expect to learn most?
1. Opinion essay (议论文)
2. Narrative writing (叙述文)
3. Expository writing (说明文)
15. Please write about your English writing learning experience.
What have you learned from the writing class in the previous years?
Do you like the previous writing course or not? Why or why not?

16. What do you expect most to learn in this class? Please give your suggestions.
(Ideas/organization/vocabulary/grammar/others)

Part Two: Overall perceptions on English writing

Answer the following questions (Question 2-7) by using the scale below to circle the response that most closely resembles your perspectives.

1: strongly disagree 2: disagree 3: Somewhat disagree
4: somewhat agree 5: agree 6: strongly agree N: no opinion

1. When I write in English, I have trouble in getting good ideas to write on.
1 2 3 4 5 6 N
2. When I write in English, I have trouble in organizing my ideas in a clear way.
1 2 3 4 5 6 N
3. When I write in English, I have trouble in finding right words to express my ideas.
1 2 3 4 5 6 N
4. When I write in English, I have trouble in using correct collocations.
1 2 3 4 5 6 N
5. When I write in English, I have trouble in using correct grammar.
1 2 3 4 5 6 N
6. When I write in English, I have trouble in using correct spelling.
1 2 3 4 5 6 N
7. When I write in English, I have trouble in using correct punctuation.
1 2 3 4 5 6 N
8. When writing in English the thing I worry about most are my ideas and the structure.
1 2 3 4 5 6 N

9. When writing in English the thing I worry about most are grammar and correctness.
1 2 3 4 5 6 N
10. When writing in English the thing I worry about most are vocabulary and correctness.
1 2 3 4 5 6 N
11. When writing in English the thing I worry about most are collocations and correctness.
1 2 3 4 5 6 N
12. When writing in English the thing I worry about most are spellings and correctness.
1 2 3 4 5 6 N
13. When writing in English the thing I worry about most are punctuations.
1 2 3 4 5 6 N

Appendix Two Questionnaire Two

Dear Students,

Thank you for answering this questionnaire. Your opinion is extremely valuable in helping me understand Chinese EFL learners' perception on the concordance exercises.

Please work at a comfortable pace and don't dwell on any one response or question. Research has shown that your initial response is usually your best response. Keep in mind that this is not a test, which means there are no right or wrong answers. Your responses should be based on the way you feel and not on how you think others would expect you to respond.

The average time to complete this questionnaire is 15 minutes.

Thank you very much!

Your Name: _____

Student ID: _____

Questionnaire Two

Evaluation on concordance exercises

Had you heard about **concordances** before you took this class? (Circle one)

上这门课之前你有没有听过语料库检索?

Yes No

Answer the following questions by using the scale below to circle the response that most closely resembles your perspectives.

1: strongly disagree 2: disagree 3: Somewhat disagree 4: somewhat agree 5: agree
6: strongly agree N: no opinion

1 The concordance exercises are helpful for me to learn the English words. 检索练习题能帮助我学习英语词汇。	1	2	3	4	5	6	N
2 The concordance exercises are helpful for me to learn the meaning of vocabulary. 检索练习题能帮助我对词汇意义的理解。	1	2	3	4	5	6	N
3 Studying concordance lines is helpful for learning the collocation of the words. 检索练习题帮助我学习词语的搭配。	1	2	3	4	5	6	N
4 Studying concordance lines is helpful for learning grammatical use of the words. 检索练习题帮助我学习词汇在句子中的用法。	1	2	3	4	5	6	N

5	Studying the concordance lines helps me memorize the usage of the target word better. 检索练习题帮助我更好的记住所学单词的具体用法。	1	2	3	4	5	6	N
6	Studying concordance lines is helpful for learning the usage of phrases. 检索练习能帮助我学习词组的具体用法。	1	2	3	4	5	6	N
7	I prefer learning the usage of the words by studying concordance lines to being taught directly by the teacher. 跟老师直接把词的用法告诉我们相比, 我比较喜欢通过自己做检索练习题来发现和学习单词的具体用法。	1	2	3	4	5	6	N
8	Studying concordance lines helps me incidentally learn more new words in the concordance output. 通过做检索练习题, 我无意中学习了一些新的单词和短语。	1	2	3	4	5	6	N
9	Studying concordance lines is helpful for my English writing on the related topic. 检索练习题对我的写作有帮助。	1	2	3	4	5	6	N
10	Studying concordance lines helps me get some expressions in English related to the writing topic. 检索练习题帮助我学习一些跟写作题目有关的词汇。	1	2	3	4	5	6	N
11	Studying the concordance lines helps me gain some ideas for my writing. 检索练习题帮助我获得一些跟写作题目相关的信息。	1	2	3	4	5	6	N
12	Learning about concordances has increased my confidence in using the words in English writing. 通过做检索练习题, 我在写作中对某些单词的用法和表达更加肯定, 从而增加了用词方面的信心。	1	2	3	4	5	6	N
13	Overall, the concordance exercises are very useful resource for my vocabulary use in English writing. 总体上, 检索练习题对我在英文写作中词汇运用方面提供了很好的学习资源。	1	2	3	4	5	6	N
14	Overall, the concordance exercises help me improve my writing quality. 总体上, 我觉得检索练习题帮助我提高写作的质量。	1	2	3	4	5	6	N
15	The concordance exercises help me memorize the key words related to the writing topic better. 检索练习题帮助我牢记一些写作关键词。	1	2	3	4	5	6	N
16	I think I can use the collocations and expressions learned from the concordance exercises in my future writing related to the similar topic. 在以后写相关题目的文章时, 我觉得我能使用在做检索练习题中所学到的词汇。	1	2	3	4	5	6	N
17	I hope we can have more concordance exercises to do in the future. 我希望以后有机会做多一些检索练习。	1	2	3	4	5	6	N

18 I have some difficulty in studying concordance lines due to time and effort spent on analyzing the data. 由于要花很多时间和精力去分析单词的用法, 我觉得做这套检索练习题有点困难。	1	2	3	4	5	6	N
19 I have some difficulty in studying concordance lines due to unfamiliar vocabulary in the data. 由于有太多的生词, 我做检索练习题是觉得有点困难。	1	2	3	4	5	6	N
20 I have some difficulty in studying concordance lines due to cut-off sentences in the exercises. 由于不完整的句子让我做这套检索练习题时感到有点困难。	1	2	3	4	5	6	N
21 I have some difficulty in studying concordance lines due to too many sentences in the exercises. 由于例句太多, 让我做这套检索练习题时感到有点困难。	1	2	3	4	5	6	N
22 I have some difficulty in studying concordance lines due to the limited number of sentences in the exercises. 由于每个词的例句太少, 让我做这套检索练习题时感到有点困难。	1	2	3	4	5	6	N
23 I have some difficulty in formulating the overall rules of the usage of the words even I spend time and effort studying the concordance lines. 通过检索练习题来总结每个词的用法我觉得有点困难, 即使我花了很多时间和精力去学习。	1	2	3	4	5	6	N
24 Overall, studying the concordance lines is time-consuming and boring. 总体上, 做检索练习题浪费时间而且很闷。	1	2	3	4	5	6	N

Appendix Three Learning Journal

Name: _____

Student ID: _____

Please answer the following question. Send your comments as an email attachment to huangzeping@gmail.com before Oct. 19, 2009, labeling your file name and email title as “your Chinese name & Student ID”.

Do you think that the vocabulary exercises helped you with your writing task on “lottery”? If yes, in what ways did it help you with your writing? If no, please explain why.

Please feel free to state your point of view. You can write either in English or Chinese.

Appendix Four Pre-writing Vocabulary Study

Pre-writing Vocabulary Study 1

Key words: objection, controversy, criticism, situation, effect

Exercise One

Direction: Study the concordance, underlining or highlighting the central group of words that stand alone, as has been done in the first example. Then answer the questions which follow. Do not worry that these are cut-off sentences—just familiarize yourself with the key words.

Objection

Study the lexico-grammatical patterns

Study the concordance lines of *objection* and answer the following questions.

1. Lottery was eventually approved. *Much of the objection to the National Lottery came from church leaders.*
2. profits to charity, but was rejected. *My personal objection to Camelot as the lottery organizer is that a large*
3. most famous businesses and Families had *a particular objection to the start of the National Lottery.*
4. All rely on participating viewers who *have no great objection to winning their moment of glory by doing their best*
5. harmless fun which won't break the bank. *The main objection to the lottery is based on the grounds that those who*
6. than that, a tax on the poor. *The main objection came from charities who predicted that charitable donate*
7. like horse-racing and casinos. *These objections became much greater with the introduction of scratch cards*
8. are how bitterly unpopular taxation is; therefore, *objections raised against the lottery concerned it being marketed*
9. or saving for a holiday or a car. *There were also objections raised to the amount of money the proposed jackpot was*
10. and psychological drawbacks. There were two main *objections against the introduction of the national lottery*
11. would have been donated to charity. *Another objection raised was that the National Lottery would*
12. The proposal of a lottery brought about *many objections and complaints.* There were, and still are, two
13. of conservative government ! *Despite the numerous objections,* the introduction of the lottery has induced a 'fever

1. Which adjectives are used before **objection(s)**? Please write down the phrases.

e.g. *the main objection*

2. Which verbs or verbs phrases are used with **objection(s)**? Please write down the phrases.

3. Which proposition commonly follows **objection(s)**?

objection(s) _____

4. Read the concordance lines carefully. Apart from the phrases of **objection**, which words or phrases or sentence structures do you think can be used in your writing task?

Controversy

Study the lexico-grammatical patterns

Study the concordance lines of *controversy* and answer the following questions.

1. most or all of this money would go to charity. *There has also been some **controversy** over* the allocation of money.
2. The statistics confirm a trend that will *reignite the **controversy** over global warming*, with the past 15 years
3. there should be a maximum jackpot of 20 million. The recent ***controversy** about* the impartiality of the head of Office,
4. almost from her beginning, the yacht has *provoked **controversy***. It was soon after Guthrie, who acquired her in the,
5. by the 90-strong Bar Council but it has *stirred **controversy*** within the Inns of Court and some traditional
6. to be a fresh avowal of rape. Since *the entire **controversy** was triggered* by those quotes, and since Depardieu
7. gained a level of legitimacy despite *myriad **controversies*** facing it .Online casinos are now competing for advertising
8. did not even exist. *The numerous **controversies** surrounding online gambling* are tied to both legal and moral issue

1. Which proposition commonly follows **controversy/controversies**?

controversy/controversies _____

2. Which adjectives are often used before **controversy/controversies**? Please write down the phrases.

3. Which verbs or verbs phrases are used with **controversy/controversies**? Please write down the phrases.

4. Read the concordance lines carefully. Apart from the phrases of **controversy/controversies**, which words or phrases or sentence structures do you think can be used in your writing task?

Criticism

Study the lexico-grammatical patterns

Study the concordance lines of *criticism* and answer the following questions.

1. that would considerably change their life. *Another criticism* is of the size of the jackpots themselves. Many people
2. slogan of the lottery 'it could be you' *came under intense criticism*. However, recent studies have shown that the vast
3. the national lottery originally claimed. *As far as the criticism directed against the government is concerned*, there was
4. say it hasn't done enough. *Egypt has already faced criticism for conducting arbitrary arrests* and indefinite
5. *Societies are facing criticism over their fees*, Caroline Merrell finds
6. *It is difficult to argue against these criticisms*. The amounts given to charity by individual donations
7. Burmese representative, Thein Tin, *rejecting the criticisms*, and defending his country's human rights record,
8. I am not going to count any. *The Minister rejected criticisms that* the Government is not doing enough to secure the

1. Which verbs or verbs phrases are used with **criticism(s)**? Please write down the phrases.

2. Which adjectives are often used before **criticism(s)**? Please write down the phrases.

3. Which propositions commonly follow **criticism(s)**? Please write down the phrases.

4. Read the concordance lines carefully. Apart from the phrases of **criticism(s)**, which words or phrases or sentence structures do you think can be used in your writing?

Situation

Study the lexico-grammatical patterns

Study the concordance lines of *situation* and answer the following questions.

1. gambling problems often *face such complicated situations* in life. It then takes several years to pay back the debt.
2. It would increase gambling and *exacerbate the situations* of many whose financial situations would not
3. professional help should be hired if *the situation gets out of hand*. Gambling casinos have an environment
4. the previous debt. *This will only worsen the situation* further. Families should be very careful
5. Many Britons *are in situations of poverty* where the lack of employment prospects
6. seen to fail the children involved by *leaving them in situations of risk*. (Secretary of State for Social Services,
7. that may change their luck. People *in desperate situations* often tend to turn to the gambling table where they test.
8. importance, for without it a car is unusable *in many situations*. Overall dimensions are another subject which is
9. have annoying traits which we demonstrate *in certain situations* and which we counterbalance with our personality
10. and tacit, intuitive, commonsense knowledge. *In some situations*, these different sources of knowledge were in
11. ability to make people laugh, even *in the most serious situations*. And many times he has shown his readiness to take

1. Which adjectives are often used before **situation(s)**? Please write down the phrases

2. Which verbs or verbs phrases are used with **situation(s)**? Please write down the phrases.

3. Which propositions are commonly used with **situation(s)**? Please write down the phrases.

4. Read the concordance lines carefully. Apart from the phrases of **situation(s)**, which words or phrases or sentence structures do you think can be used in your writing?

Effect

Study the lexico-grammatical patterns

Study the concordance lines of *effect* and answer the following questions.

1. a profit on it . Many have also *speculated as to the effects* of being a jackpot winner. *Sudden riches overnight* see
2. term implications on gambling habits plus the *positive effects* will have to be considered before a decision on its
3. in the family needs to be *educated on the harmful effects* of gambling too much. Put security locks on the gamble
4. a print before exposing it to *get even more noticeable effects*, but there's no solution other than to reprint the
5. the Georgian capital, Tbilisi. He *compares the terrible effects* of multiple abortions in Georgia to the problem of
6. Thus, previewing served here *to predict the potential effects* of Deirdre's illness and of her personality traits on
7. The oil spill didn't seem to *have any detrimental effects* on the salmon run, so we didn't observe any
8. because it has *had a significant effect* on the economy. As well as causing a fall in charity
9. church have expressed great concern at the *effect on low income families* who spend more than they can
10. 'flutter' on the National lottery each week *has no effect* on their financial position. This kind of people can a

1. Which adjectives are used before **effect(s)**? Please write down the phrases.

2. Which verbs or verbs phrases are used with **effect(s)**? Please write down the phrases.

3. Which prepositions are used with **effect(s)**?

4. Read the concordance lines carefully. Apart from the phrases of **effect**, which words or phrases do you think can be used in your writing?

Pre-writing Vocabulary Study 2

Key words: misery, charity, addiction, casino, jackpot

Exercise Two

Direction: Study the concordance, underlining or highlighting the central group of words that stand alone, as has been done in the first example. Then answer the questions which follow. Do not worry that these are cut-off sentences—just familiarize yourself with the key words.

Charity

Study the lexico-grammatical patterns

Study the concordance lines of *charity* and answer the following questions.

1. *It was argued that the amount of money going to **charity** had been emphasized too much and the figures exaggerate*
2. *exaggerated. Thus, it was argued, donations to **charity** would be heavily hit by the introduction of the lottery*
3. *argue against these criticisms. The amounts given to **charity** by individual donations have been hit as heavily as the*
4. *any believe far more lottery-generated revenue goes to **charity** than it actually does. In response, it can only be sa*
5. *because they would think they had been giving money to **charity**. It is now known that only eight pence in the pound*
6. *ated by its funds, it provides an important source of **charity** which wouldn't otherwise be true. Beneficiaries have*
7. *national lottery was the issue of raising funds for home **charities** and needy causes. This issue is coupled with*
8. *one main advantage to the lottery. It has helped **charities** and societies around the country to achieve*

1. which nouns or noun phrases does **charity/charities** often collate? Please write down the phrases.

2. Read the concordance lines carefully. Apart from the phrases of **addiction**, which words or phrases or sentence structures do you think can be used in your writing?

Addiction

Study the lexico-grammatical patterns

Study the concordance lines of *addiction* and answer the following questions.

1. it was argued , *posed a much more real danger of* **addiction** because the results could be determined
2. as a bad thing which would tend to *become a hopeless* **addiction**. These people tended to be those whose ideas
3. problem gamblers it is *a dangerous* **addiction**. This book looks at current UK gambling trends
4. in profits. *There were also problems raised about* **addiction** and people younger than sixteen obtaining numbers.
5. future riches. For many this is not the case, and **addiction** *means large debts and results in violence* and stealing
6. to obtain extra money; *It was argued that this* **addiction** *is on a par with drug addiction* and just as hard to
7. with a negative expected return. *Gambling, an* **addiction** *for some people, can result into a huge disaster* if
8. it may not harm you . *It is only when it becomes an* **addiction** *that it can strip you of not just your money* but also
9. Participating in lotteries can also be *addictive* and these **addictions** *should be controlled* as good money is used to
10. gambling, they're `investing.'" *Those with gambling* **addictions** *stand to lose a lot more than money:* 38 percent
11. perceived to be much *more serious mental health and* **addiction** *problems*. Doctors said the national lottery had
12. Because there are a lot of things written *about drug* **addictions** *or alcohol addictions* that type of thing but
13. Love melts away tumors, *cures* **addictions**, banishes fear, catalyzes miracles, transform live

Part Two: Study the lexico-grammatical patterns

Study the concordance lines carefully and answer the following questions.

1. Which adjectives are often used before **addiction(s)**? Please write down the phrases.

2. Which nouns or nouns phrases are used before **addiction(s)**? Please write down the phrases.

3. Which verbs or verbs phrases are used with **addiction(s)**? Please write down the phrases.

4. Which prepositional phrases are used with **addiction(s)**? Please write down the phrases.

5. Read the concordance lines carefully. Apart from the phrases of **addiction**, which words or phrases or sentence structures do you think can be used in your writing?

Misery

Study the lexico-grammatical patterns

Study the concordance lines of *cause* and answer the following questions.

1. on the whole, *it has caused more misery than happiness*. Abolition may be too much to hope for
2. The churches' main argument is that *it is causing misery to thousands of people*. Those who have become
3. next weekend it will *bring joy to many and misery* to just a few. When the national lottery was first
4. failures. 3000 jobs lost and 150 families *suffering misery of having their homes repossessed*. If efforts are
5. the loss of a lover could *be responsible for such misery*. Top model Helena Christensen sits alone at a
6. I didn't *want to add to your miseries*, Russell. I didn't want to burden you. But I am scared
7. often unimaginable horrors, *impossible personal miseries*, the squalor of defeat and loneliness
8. the one to come after my death. I *am undergoing the miseries of this life* because of the crimes of somebody else
9. than it was. The present trend for *exaggerating the miseries* of run-down estate Britain
10. that the navy *was ultimately responsible for the miseries*, both internal and external, that now beset Germany.
11. suffering is not relative; we *all feel our miseries and hardships absolutely*. I balk at the count-your-

Part Two: Study the lexico-grammatical patterns

Study the concordance lines carefully and answer the following questions.

1. Which adjectives are often used before **misery/miseries**? Please write down the phrases.

Which verbs or verbs phrases are used with **misery/miseries**? Please write down the phrases.

3. Which prepositional phrase is used with **misery/miseries**? Please write down the phrases.

4. Read the concordance lines carefully. Apart from the phrases of **misery**, which words or phrases or sentence structures do you think can be used in your writing?

Jackpot

Study the concordance lines of **jackpot** and choose its meaning

a.	A kind of container of liquor of other drink
b.	A top prize or reward in a competition or lottery

1. Firstly, the amount of money which the **jackpot winners** were likely to receive was deemed as 'too much'
2. press until they tracked him down . *Arguments for **jackpot to be capped** have been put forward* by the opposition in
3. retained, but not in its present form. I think that **jackpots should be capped** at 2 million pounds, and the prize
4. The frenzy seen just the other week, as *the estimated **jackpot rose to** over 20 million*, was quite startling. It is a
5. The probability of *winning the **jackpot*** is approximately 1 in 14 million and the probability of
6. Many people claim, *after winning a **jackpot**, you slowly lose friends and can become completely cut off*
7. the risk that your six numbers might *hit the **jackpot** but with odds of 14 millions to one against*, I

8. Many have also speculated as to *the effects of being a **jackpot** winner. Sudden riches overnight* seems like a dream
9. of the lucky winners have said themselves that the **jackpot** *had ruined their life, alienating them from friends and*
10. *winning the Lottery **jackpot** can often ruin peoples lives. Many people claim*
11. *illusion of mystery and excitement by increasing the **jackpot** to 40 million and introducing instant scratches.*
12. *It has also been alleged that the **jackpots** are too high* most of the lucky winners have said

Part Two: Study the lexico-grammatical patterns

Study the concordance lines carefully and answer the following questions.

1. Which nouns or nouns phrases are used before **jackpot**? Please write down the phrases.

2. Which nouns or nouns phrases are used after **jackpot**?

3. Which adjectives are used before **jackpot**?

4. Which verbs or verbs phrases are used with **jackpot**?

5. Read the concordance lines carefully. Apart from the phrases of **jackpot**, which words or phrases do you think can be used in your writing?

Casino

Study the lexico-grammatical patterns

Study the concordance lines of *casino* carefully and answer the following questions

1. In Moscow alone, there were more than 500 *legal casinos and gaming halls*. Police have been patrolling the streets
2. in order to ensure the doors are firmly shut. *The spread of casinos has provoked distaste* among many Russians, especially
3. those who can now enjoy whenever they want *at online casinos*. Some credit-card issuers, recognising the lure of online
4. gaming, have other options in *paying for bets placed at online casinos*, including debit cards, electronic-funds transfer from bank
5. accounts and *tighter regulations on the territory's casinos*. It is as if the gold is running out in the Klondike..
6. There has been a proliferation of *casinos* and marketing of lotteries, and a proliferation of instant scratch-off tickets
7. sold to more than two million Americans *regularly place bets at casino websites*, despite the fact that none are based within
8. the territory. *Gambling casinos* have an environment that attracts a person and engages him totally into the energetic buzz of players
9. who play on all sorts of gambling games that you *get to play at casinos* and you have to pay by credit card. Poker and jackpot
10. are the most popular. However, the most famous form of the game remains *casino gambling*. Las Vegas is very famous for its casinos that
11. offer a wide range of entertainment. Las Vegas is very famous for *its casinos* that offer not just gambling avenues but also other opportunities for entertainment

1. Which verbs or verbs phrases are used with **casino(s)**? Please write down the phrases.

2. Which adjectives are used before **casino(s)**? Please write down the phrases.

3. Which nouns or nouns phrases are used with **casino(s)**? Please write down the phrases.

4. Which preposition is used with **casino(s)**? Please write down the phrases.

5. Read the concordance lines carefully. Apart from the phrases of **casino(s)**, which words or phrases do you think can be used in your writing?

CUHK Libraries



004779146