
Theses and Dissertations

2006

Effects of using corpora and online reference tools on foreign language writing: a study of Korean learners of English as a second language

Kyosung Koo
University of Iowa

Copyright 2006 Kyosung Koo

This dissertation is available at Iowa Research Online: <http://ir.uiowa.edu/etd/65>

Recommended Citation

Koo, Kyosung. "Effects of using corpora and online reference tools on foreign language writing: a study of Korean learners of English as a second language." PhD (Doctor of Philosophy) thesis, University of Iowa, 2006.
<http://ir.uiowa.edu/etd/65>.

Follow this and additional works at: <http://ir.uiowa.edu/etd>



Part of the [First and Second Language Acquisition Commons](#)

EFFECTS OF USING CORPORA AND ONLINE REFERENCE TOOLS ON
FOREIGN LANGUAGE WRITING: A STUDY OF KOREAN LEARNERS OF
ENGLISH AS A SECOND LANGUAGE

by
Kyosung Koo

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Second Language Acquisition
in the Graduate College of
The University of Iowa

May 2006

Thesis Supervisor: Associate Professor James P. Pusack

ABSTRACT

The general aim of this study is to better understand aspects of using reference tools for writing and to identify technologies that can assist foreign language writers. The specific purpose of this study is to look closely at how English as a Second Language (ESL) students from Korea use a corpus as a reference tool in conjunction with dictionaries when paraphrasing English newspaper articles. The participants were Korean graduate students with advanced English proficiency (N=10). Their task was to paraphrase an English newspaper article.

The results show that purposes for using a concordancing program include collocations, definitions, context, and parts of speech. The subjects looked for a variety of information in a concordancing program, including prepositions, authentic samples, and the context in which the search terms were used. Reasons for using dictionaries include definitions, parts of speech, and sample sentences. The most common strategy was to combine reference tools, while the second most common was to use a specific search word. Subjects who used more than one tool for a search or performed multiple searches were more successful in finding what they were looking for.

A concordancing program enabled users to see multiple examples of everyday language use. By using the concordancing program, learners were able to see words that were used most frequently, their patterns, and collocations. Learners took more responsibility for their language learning, as they became researchers in their own right. They gained confidence as L2 writers as they had inside access to linguistic resources. The subjects became more independent and were able to solve their own writing and linguistic problems as they became more aware through the use of authentic texts. In this study, the subjects found the corpora to be useful for sentence-level composition and revision. Overall, the use of reference tools led to an improvement in the accuracy of

writing. A concordancing program played an important role in defining the structure and context of English phrases and sentences.

Abstract Approved: _____
Thesis Supervisor

Title and Department

Date

EFFECTS OF USING CORPORA AND ONLINE REFERENCE TOOLS ON
FOREIGN LANGUAGE WRITING: A STUDY OF KOREAN LEARNERS OF
ENGLISH AS A SECOND LANGUAGE

by
Kyosung Koo

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Second Language Acquisition
in the Graduate College of
The University of Iowa

May 2006

Thesis Supervisor: Associate Professor James P. Pusack

Copyright by
KYOSUNG KOO
2006
All Rights Reserved

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Kyosung Koo

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Second Language Acquisition at the May 2006 graduation.

Thesis Committee: _____
James P. Pusack, Thesis Supervisor

Stephen M. Alessi

L. Kathy Heilenman

Sue K. Otto

Carol J. Severino

To my wife, Beth, whom I love.

What the concordancer does is make the invisible visible.

Chris Tribble
Concordancing and an EAP writing programme

ACKNOWLEDGMENTS

I would like to take this opportunity to thank those who so generously gave of their time, expertise, and support in order to help me successfully complete this dissertation. Professor James Pusack was a tremendous help throughout the entire process. He has been my academic advisor during my time at the University of Iowa and a great inspiration. Also, Professor Carol Severino, Director of the Writing Center at the University of Iowa, assisted in recruiting the research subjects. I want to thank Professor Stephen Alessi for his input on research design and methodology. Professor Kathy Heilenman's expertise in corpus and analysis of data also contributed to the successful completion of this project. Sue Otto, Professor and Director of the Language Media Center at the University of Iowa advised and supported me in the field of instructional technology, and was a great source of encouragement to me to complete this dissertation when I moved away from Iowa. Finally, I would also like to thank Joshua Thoms for his contributions to this research during his busy schedule.

ABSTRACT

The general aim of this study is to better understand aspects of using reference tools for writing and to identify technologies that can assist foreign language writers. The specific purpose of this study is to look closely at how English as a Second Language (ESL) students from Korea use a corpus as a reference tool in conjunction with dictionaries when paraphrasing English newspaper articles. The participants were Korean graduate students with advanced English proficiency (N=10). Their task was to paraphrase an English newspaper article.

The results show that purposes for using a concordancing program include collocations, definitions, context, and parts of speech. The subjects looked for a variety of information in a concordancing program, including prepositions, authentic samples, and the context in which the search terms were used. Reasons for using dictionaries include definitions, parts of speech, and sample sentences. The most common strategy was to combine reference tools, while the second most common was to use a specific search word. Subjects who used more than one tool for a search or performed multiple searches were more successful in finding what they were looking for.

A concordancing program enabled users to see multiple examples of everyday language use. By using the concordancing program, learners were able to see words that were used most frequently, their patterns, and collocations. Learners took more responsibility for their language learning, as they became researchers in their own right. They gained confidence as L2 writers as they had inside access to linguistic resources. The subjects became more independent and were able to solve their own writing and linguistic problems as they became more aware through the use of authentic texts. In this study, the subjects found the corpora to be useful for sentence-level composition and revision. Overall, the use of reference tools led to an improvement in the accuracy of

writing. A concordancing program played an important role in defining the structure and context of English phrases and sentences.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES.....	xii
CHAPTER I. INTRODUCTION.....	1
1.1. Background.....	1
1.2. The role of corpora in language teaching and linguistics.....	1
1.3. Dictionaries.....	2
1.4. Corpora and SLA.....	3
1.5. Overview of the study	4
1.6. Research questions.....	5
CHAPTER II. REVIEW OF THE LITERATURE.....	7
2.1. Introduction	7
2.2. Corpora and concordancing.....	7
2.2.1. Corpora.....	7
2.2.2. Concordancing	14
2.3. Dictionary use.....	20
2.3.1. Dictionary use and benefits.....	20
2.3.2. Concerns about dictionary use	27
2.4. Corpora in second language acquisition.....	34
2.4.1. Context.....	35
2.4.2. Collocation.....	41
2.4.3. Active and discovery-style learning	43
2.4.4. Strategies and techniques of corpus use in second language learning and teaching.....	46
2.4.5. Using parallel corpora for second language learning	50
2.5. Concerns about concordancing.....	52
CHAPTER III. METHODOLOGY.....	59
3.1. Data gathering techniques	59
3.2. Pilot study	64
3.2.1. Research questions for the pilot study.....	64
3.2.2. Subjects.....	64
3.2.3. Materials	65
3.2.4. Procedures.....	66
3.2.5. Results	68
3.3. Main study	71
3.3.1. Subjects.....	71
3.3.2. Materials	72
3.3.3. Procedures.....	75
3.4. Analysis	79
3.4.1. Interrater reliability	79
3.4.2. Units of analysis.....	81
3.4.3. Coding	82

CHAPTER IV. RESULTS.....	87
4.1. What are the reasons why ESL learners consult a concordancing program while writing in English?	87
4.2. What kind of information do learners look for in a corpus and dictionaries while writing in English?	97
4.3. What strategies do learners use when they consult a corpus?	105
4.4. How does a corpus complement the use of dictionaries?.....	116
4.4.1. Combination of thesaurus and <i>MonoConc</i>	117
4.4.2. Combination of English-Korean dictionary and <i>MonoConc</i>	121
4.4.3. Combination of English-English dictionary and <i>MonoConc</i>	135
4.4.4. Combination of <i>MonoConc</i> and <i>MonoConc</i>	137
4.5. What are the effects of using a corpus on the accuracy of learners' writing?	143
4.5.1. Repeated search word.....	147
4.5.1.1. Repeated search word: "blame".....	147
4.5.1.2. Repeated search word: "criticize".....	153
4.5.1.3. Repeated search word: "accuse".....	157
4.5.2. <i>MonoConc</i> leading to wrong grammar.....	163
4.5.3. Using phrases as search terms.....	169
4.6. What are learners' attitudes toward using a corpus as a reference tool?.....	172
CHAPTER V. DISCUSSION.....	176
5.1. Purposes of using a concordancing program.....	176
5.2. Learner strategies for using a concordancing program.....	181
5.3. Corpus use and accuracy of writing.....	186
5.4. Learners' attitudes toward using a corpus as a reference tool.....	189
5.5. Implication for corpus use in language teaching	191
5.6. Future research.....	195
APPENDIX A. NEWSPAPER ARTICLE ON KOREAN PRESIDENT IMPEACHMENT	197
APPENDIX B. NEWSPAPER ARTICLE ON NORTH KOREAN NUCLEAR PROGRAM.....	201
APPENDIX C. PLANS FOR PARAPHRASING SESSIONS	205
APPENDIX D. DIRECTIONS FOR PARAPHRASING	207
APPENDIX E. DIRECTIONS FOR STIMULATED RECALL SESSION	209
APPENDIX F. DEMOGRAPHIC DATA.....	211
APPENDIX G. SURVEY.....	214
APPENDIX H. IRB DOCUMENTS: INFORMED CONSENT DOCUMENTS	218
REFERENCES.....	224

LIST OF TABLES

Table

3.1. Individual subjects' use of reference tools when paraphrasing and summarizing.....	69
3.2. Use of reference tools for different tasks.....	69
3.3. Use of reference tools for different methods	71
3.4. Time frame of the meetings	79
4.1. Subject JWK using <i>MonoConc</i> with search term "accus"	89
4.2. Subject HJS using <i>MonoConc</i> with search term "destruction @ country"	90
4.3. Subject JWK using <i>MonoConc</i> with search term "restrain"	91
4.4. Subject MKS using <i>MonoConc</i> with search term "planned visit"	92
4.5. Subject HJS using <i>MonoConc</i> with search term "decline"	92
4.6. Subject MAP using thesaurus with search term "rejection"	93
4.7. Subject KWA using English-Korean dictionary with search term "inconclusive"	94
4.8. Subject KYL using Korean-English dictionary with search term "dedam"	95
4.9. Subject YJK using English-English dictionary with search term "noose" and "accord"	95
4.10. Subject KYL using English-English dictionary with search term "bottom-line".....	96
4.11. Subject JYP using <i>MonoConc</i> with search term "skeptical about".....	100
4.12. Subject JDO using <i>MonoConc</i> with search term "agreed"	101
4.13. Subject MKS using thesaurus with search term "dismantle".....	102
4.14. Subject MAP using English-Korean dictionary with search term "move"	103
4.15. Subject KWA using Korean-English with search term "palsang"	103
4.16. Subject JWK using English-English dictionary with search term "explicitly"	104
4.17. Subject HJS using thesaurus with search term "reject" and <i>MonoConc</i> with search term "decline"	107
4.18. Subject HJS using <i>MonoConc</i> with search term "declin"	108

4.19. Subject JDO using <i>MonoConc</i> with search term "visit"	109
4.20. Subject JWK using thesaurus with search terms "carry" and “monitory”	110
4.21. Subject KWA using <i>MonoConc</i> with search term "insist"	111
4.22. Subject KWA <i>MonoConc</i> with search term "press"	112
4.23. Subject KWA using <i>MonoConc</i> with search term "reproach"	113
4.24. Subject KWA using <i>MonoConc</i> with search term "make an agreement"	113
4.25. Subject JYP using <i>MonoConc</i> with search term "announc*"	114
4.26. Subject JYP using <i>MonoConc</i> with search term "blame*"	115
4.27. Subject HJS using thesaurus with search term “reject” and <i>MonoConc</i> with search term "decline"	118
4.28. Subject MKS using thesaurus with search term “inconclusive” and <i>MonoConc</i> with search term "ambiguity"	120
4.29. Subject JYP using English-Korean dictionary with searcher term “reject” and <i>MonoConc</i> with search term "reject*"	121
4.30. Subject MAP using English-Korean dictionary with search term “trust” <i>MonoConc</i> with search term "trust"	123
4.31. Subject MKS using English-Korean dictionary with searcher term “boon” and <i>MonoConc</i> with search term "boon"	124
4.32. Subject JDO using Korean-English dictionary with search term “chujanghada” and <i>MonoConc</i> with search term "assert"	125
4.33. Subject KWA using Korean-English dictionary with search term “tae-ung” and <i>MonoConc</i> with search term "respond*"	126
4.34. Subject MKS using Korean-English dictionary with search term “hasu” and <i>MonoConc</i> with search term "underling"	128
4.35. Subject MKS using Korean –English dictionary with search term “nagada” and <i>MonoConc</i> with search term "step forward"	129
4.36. Subject JHC using <i>MonoConc</i> with search term "plot" and English-Korean dictionary with search term “plot”	131
4.37. Subject KWA using <i>MonoConc</i> with search term "criticiz*” and Korean- English dictionary with search term “binan”	133
4.38. Subject MKS using <i>MonoConc</i> with search term "fitful" and English-Korean dictionary with search term “fitful”	134
4.39. Subject JWK using <i>MonoConc</i> with search term "put@forward" and English- English with search term “put forward”	136

4.40. Subject KYL using <i>MonoConc</i> with search term "aim"	139
4.41. Subject JYP using <i>MonoConc</i> with search term "announc*"	140
4.42. Subject JYP using <i>MonoConc</i> with search term "skeptical"	141
4.43. Subject HJS using <i>MonoConc</i> with search term "destruction"	143
4.44. Reference Tool Use and Grammaticality.....	145
4.45. Subject JDO using <i>MonoConc</i> with search term "blame"	148
4.46. Subject JYP using <i>MonoConc</i> with search term "blame* that"	149
4.47. Subject KWA using <i>MonoConc</i> with search term "blame"	150
4.48. Subject KWA using thesaurus with search term "blame" and <i>MonoConc</i> with search term "condemn"	152
4.49. Subject JHC using <i>MonoConc</i> with search term "criticize*"	154
4.50. Subject JYP using <i>MonoConc</i> with search term "criticism"	155
4.51. Subject KWA using <i>MonoConc</i> with search term "criticiz*"	156
4.52. Subject JDO using <i>MonoConc</i> with search term "accused"	158
4.53. Subject JHC using <i>MonoConc</i> with search term "accuse*"	159
4.54. Subject JYP using <i>MonoConc</i> with search term "accus* @ of"	160
4.55. Subject YJK using <i>MonoConc</i> with search term "accus*"	162
4.56. Subject JDO using <i>MonoConc</i> with search term "as usually"	164
4.57. Subject JHC using <i>MonoConc</i> with search term "settle"	165
4.58. Subject JWK using thesaurus with search term "supervise" and <i>MonoConc</i> with search term "watch over"	167
4.59. Subject JYP using <i>MonoConc</i> with search term "conflict over"	169
4.60. Subject JYP using <i>MonoConc</i> with search term "made it clear that"	170
4.61. Subject YJK using <i>MonoConc</i> with search term "negotiating table"	171
4.62. Survey questions and results	174

LIST OF FIGURES

Figure

2.1. The Brown Corpus: sample text without tags.....	9
2.2. The Brown Corpus: sample text with tags.....	9
2.3. Concordancing program: <i>MonoConc Pro 2.2</i> (Barlow, 2002) and <i>English Gigaword</i> (Graff, 2003).....	15
2.4. Concordancing program search results: <i>MonoConc Pro 2.2</i> (Barlow, 2002) and <i>English Gigaword</i> (Graff, 2003).....	15
2.5. Concordancing program collocation frequency: <i>MonoConc Pro 2.2</i> (Barlow, 2002) and <i>English Gigaword</i> (Graff, 2003).....	16
2.6. <i>ParaConc</i> (Search word “head”) (Barlow, 2001).....	18
2.7. Concordancing program search results in KWIC format: <i>MonoConc Pro 2.2</i> (Barlow, 2002) and <i>English Gigaword</i> (Graff, 2003).....	18
3.1. Use of corpus and other reference tools	70
3.2. Collocation frequency statistics (search word: work*)	73
3.3. <i>Yahoo</i> Online Korean Dictionaries.....	75
3.4. Screen recording for stimulated recall.....	78
3.5. Units of analysis	82
4.1. Reasons for using <i>MonoConc</i>	88
4.2. Types of search word.....	98
4.3. Information looked for in <i>MonoConc</i>	99
4.4. Strategies for using a concordancing program.....	106
4.5. Ratio of Grammaticality (Correct: Incorrect: Null).....	145
4.6. Ratio of Grammaticality (Correct: Incorrect)	146
4.7. Ratio of <i>MonoConc</i> Use and Grammaticality.....	146
4.8. Survey results.....	175

CHAPTER I

INTRODUCTION

1.1. Background

A corpus is a collection of authentic language production. The size of a corpus can vary from thousands to millions of words. So, why, one may ask, is such a large collection of language necessary? From Chomsky's viewpoint (McEnery & Wilson, 2001), our knowledge of rules enables us to create original sentences. His theory is that by nature, corpora are incomplete and skewed. Some sentences are found in corpora because they are commonly used constructions, while others may be found in corpora by chance. An empirical approach can be carried out by observing natural language data through a corpus. For example, how is it possible to know why "which" is used instead of "that" in a relative clause? For this question, Chomsky may say that he knows because he is a native speaker of English. Conversely, a corpus linguist would say to look in the corpus and find out.

Corpora serve many purposes. Historically, they were used in the production of dictionaries and language textbooks. For this purpose, a corpus must be tagged, or given annotations such as the part of speech and syntactical placement in a sentence. In order to analyze such a large quantity of data, the text must be machine-readable and loaded into a concordancing program. A corpus can also be a source of empirical (quantitative) data for linguistic research. Such research may include speech, lexical studies, grammar, semantics, pragmatics, discourse analysis, and the teaching of language and linguistics.

1.2. The role of corpora in language teaching and linguistics

This study focuses on using a corpus for the purpose of language learning. Kennedy (1987a; G. Kennedy, 1987b), for example, has researched how to express

quantification and frequency in English as a Second Language (ESL) textbooks. Foreign or second language teaching includes a category entitled English for Specific Purposes, or ESP — for example, teaching medical English to medical students or legal English terminology to law students. In cases such as these, having access to a corpus with this specific information can be invaluable to students. Computer-assisted language learning (CALL) is another application of corpora in language and linguistics in which a corpus, or corpora, may be utilized. Lancaster University has been researching the role computer software that is corpus-based plays with grammatical analysis taught to undergraduates (McEnery & Wilson, 1993). Multilingual parallel corpora are also becoming more available and serve as a basis for teaching translation. These corpora provide side-by-side examples of style and idioms in multiple languages and can create exercises that students can use to compare their own translations with an actual professional translation (Zanettin, 1994).

1.3. Dictionaries

As mentioned above, corpora have historically been utilized in creating dictionaries. Dictionaries have traditionally been the main reference tool used by language learners, but are largely deficient in the more holistic information included in corpora. In order to understand the advantages of using corpora, it is important first to recall what is offered by dictionaries. They provide definitions, parts of speech, phonetic spelling, sample sentences, related phrases, and word etymology. Different types of dictionaries are available, such as monolingual, bilingual, and bilingualized (see definitions in Section III). Both language professionals and language learners have differing opinions about which is the best.

Some studies show that dictionaries do not provide enough context for synonyms, which may result in incorrect choices of words (Harvey & Yuill, 1997). Sometimes the format of dictionaries may confuse the users (Parry, 1991, 1993); as a result, learners

may not make use of all of the features, such as grammar-related information (Bejoint, 1981) and make incorrect lexical choices (Varantola, 1998). For these reasons, language learners need strong dictionary skills to understand unfamiliar words. Barlow (Barlow, 1996b) claims that dictionaries do not have enough context to be of any value to language learners. Parallel texts, he suggests, may be richer and more flexible, providing patterns that language learners can use as a search tool.

1.4. Corpora and SLA

By using corpora in conjunction with dictionaries, students may reap more benefits by looking at authentic examples, context, and word collocations. Additionally, language learners who have access to corpora experience a more student-centered and active learning environment. The following is an explanation of each of these benefits.

By using corpora in a second language or foreign language class, students will be exposed to authentic examples instead of fixed or artificial examples not necessarily used in daily life. Use of fabricated examples may distort learners' views of actual language use (Flowerdew, 1993). Most sources of corpora are collected from newspapers, magazines, novels, actual conversations, and the like. Learners are provided with a wealth of examples taken from these sources.

Another advantage that a corpus brings to second language acquisition is the context it provides in the examples. By looking at examples, learners can understand the context in which words should be used. Learners may be able to discover the meaning of a word by inferencing (Stevens, 1991b). However, there may be some contradictory claims that context is not helpful for inferring vocabulary meaning through context (Maddalena, 2001). If the context is associated with the search word, there could still be some benefits (Laufer & Shmueli, 1997). Context also becomes an issue when learners' proficiency is too low to infer the meaning of words from context. In this case, using a parallel corpus could be a solution (St. John, 2001).

Corpora can also create an environment in which inductive learning takes place. This allows for students to be in control of their learning (Barlow, 1996b; C. Kennedy & Miceli, 2001). In this environment, the role of the students becomes that of linguistic researchers in which they explore the data and create their own rules and conclusions, which also changes the dynamics of teacher-student interactions. Johns (1991a) dubbed this the Data-Driven Learning (DDL) approach. The methodology of this approach allows learners to select strategies that allow them to perceive similarities and differences of word items.

When providing examples, corpora also simultaneously display word collocations via the concordancing program into which they have been loaded. Collocations are words that commonly appear together in a sentence. By looking at collocational frequencies, learners can see preceding and following information for their search word. For learners, this kind of information is also a reliable indicator of correct grammar structures. Concordancing programs are organized in such a way that language learners can understand lexical items and their relationship to each other. While this is not the case when using a dictionary, a dictionary still provides a small number of collocations.

1.5. Overview of the study

The general aim of this study is to better understand aspects of using reference tools for writing and to identify technologies that can assist foreign language writers. The specific purpose of this study is to look closely at how English as a Second Language (ESL) students from Korea use a corpus as a reference tool in conjunction with dictionaries when paraphrasing English newspaper articles.

The participants were Korean graduate students with advanced English proficiency (N=10). Their task was to paraphrase an English newspaper article. This task was chosen based on the results of the pilot study, which was conducted prior to the actual study. The pilot study provided an opportunity to test two data-gathering

techniques (think-aloud and stimulated recall protocols). Further details about the pilot study are presented in Chapter III, Methodology. The corpus they consulted was *English Gigaword*, published by Linguistic Data Consortium. *MonoConc Pro* was used as a concordancing program. The participants were allowed to use the online *Yahoo* Korean dictionaries, which include Korean-English, English-Korean, English-English, and a thesaurus.

The results of this research study will be beneficial for designing ESL writing classes utilizing the data-driven learning approach. Data-driven learning places an emphasis on the exploration of authentic materials and authentic exploratory tasks in a student-centered environment as opposed to traditional drill exercises.

The results show how beneficial consulting a corpus is in ESL writing, and since the students are the primary beneficiaries of using such reference tools, ESL educators can adapt the results to create more effective and meaningful learning environments. Students will then be exposed to authentic discourse that, as a number of corpus linguists point out, is generally the most useful for learners. By being exposed to authentic texts, students can expand their understanding of word functions in particular contexts. Students will also be able to make inductive discoveries about English while they are examining a number of examples of vocabulary items in context. This will result in student-centered learning that promotes self-confidence and mastery of the learning process.

1.6. Research questions

This study will extend the horizon of corpus linguistics research. Most corpus linguistic studies have used corpora as data and concordancing programs as investigation tools for language use. By contrast, this research study focuses on using corpora and concordancing programs as learning tools. Since writing activities are text-oriented and make use of word combinations and lexical patterns, a corpus approach clearly has the

potential to benefit second language writing. This research study should give corpus linguists a new perspective on the use of corpora. The foci of this research are to better understand aspects of using reference tools for writing and to identify technologies that can assist foreign language writers. The study will encompass the following six questions:

- (1) What are the reasons that ESL learners consult a concordancing program while writing in English?
- (2) What kind of information do learners look for in a corpus and dictionary while writing in English?
- (3) What strategies do learners use when they consult a corpus?
- (4) How does a corpus complement the use of dictionaries?
- (5) What are the effects of using a corpus on the accuracy of learners' writing?
- (6) What are learners' attitudes toward using a corpus as a reference tool?

CHAPTER II

REVIEW OF THE LITERATURE

2.1. Introduction

The growing availability of innovative technology has allowed corpora to be used more frequently as a reference tool for language teachers and learners. Traditionally, dictionaries have been used as the primary reference tool in second language classrooms. Online dictionaries provide definitions of words and phrases that include some context and grammar. However, learners sometimes become confused about which definitions are the ones they are searching for in the context in which they are used. Moreover, learners do not always receive sufficient information about how to use the word or phrase in an original sentence. Concordancing programs, on the other hand, are structured to solve these problems by providing authentic examples in realistic contexts. Through the provided context, the learner can also learn about grammatical structures by examining a variety of examples. In addition, concordancing programs provide an active environment in which learners become researchers and test their hypotheses. They encounter problems along the way and revise their hypotheses by having direct access to the data. In the following sections, I will discuss the definition of a corpus, the design of concordancing programs, ways in which a corpus complements dictionary use, how to use a corpus for the purposes of teaching and learning, and suggestions for using a corpus.

2.2. Corpora and concordancing

2.2.1. Corpora

The word “corpus” originates from the Latin word meaning “body.” In modern linguistic terms, a corpus is a large collection of language production that can be used to investigate lexis, syntax, text and discourse, regional differences, differences between

learners and native speakers, and historical changes. In broad terms, corpus linguistics is the study of linguistic phenomena by means of a set of investigative tools such as corpora and concordancing programs. Its main features include computational techniques and a large collection of language samples in the form of written or transcribed spoken language. For example, a corpus may consist of texts taken from newspapers, journals, books, or other speech productions. The texts can then be analyzed by a concordancing program and immediately displayed on a monitor.

In order to analyze a corpus, the texts must be in machine-readable format. This means that a linguist can use a computer to call up a number of examples of a word or phrase in context from millions of words in a short time. With this text format and computer technology, not only can dictionaries be produced more quickly than before, but the information can also be up to date and more precise. A computer-readable corpus may consist of raw texts that are taken directly from the sources without any additional information added. For concordancing purposes on a computer, many corpora additionally contain some types of linguistic information. This linguistic information is called a “markup,” “annotation,” or “tag.” Depending on the purposes of concordancing, this information can be grammatical, prosodic, semantic or historical. Grammatical annotation is the most common type. In a grammatically annotated corpus, each word is assigned a word class label, which is a tag that indicates the part of speech. The Brown Corpus, the Lancaster-Oslo/Bergen Corpus, and the British National Corpus are examples of grammatically annotated corpora. Figure 2.1 and Figure 2.2 show a small portion of two versions of the Brown Corpus: without and with tags. Figure 2.1 contains the original or raw text and line numbers, while Figure 2.2 contains grammatical tags as well as the original text and line numbers. Each tag represents grammatical functions of the tagged word. For example, “AT” in the first line of SA01:1 explains that “the” is an article, “NP” shows that “Fulton” is a proper noun, “VBD” indicates that “said” is a past-tense verb.

Figure 2.1. The Brown Corpus: sample text without tags

A01 0010 The Fulton County Grand Jury said Friday an investigation
 A01 0020 of Atlanta's recent primary election produced "no evidence" that
 A01 0030 any irregularities took place. The jury further said in term-end
 A01 0040 presentments that the City Executive Committee, which had over-all
 A01 0050 charge of the election, "deserves the praise and thanks of the
 A01 0060 City of Atlanta" for the manner in which the election was conducted.
 A01 0070 The September-October term jury had been charged by Fulton
 A01 0080 Superior Court Judge Durwood Pye to investigate reports of possible
 A01 0090 "irregularities" in the hard-fought primary which was won by
 A01 0100 Mayor-nominate Ivan Allen Jr&.

Figure 2.2. The Brown Corpus: sample text with tags

SA01:1 the_AT Fulton_NP County_NN Grand_JJ Jury_NN said_VBD Friday_NR
 an_AT investigation_NN of_IN Atlanta's_NP\$ recent_JJ primary_NN election_NN
 produced_VBD no_AT evidence_NN that_CS any_DTI irregularities_NNS
 took_VBD place_NN ._.

SA01:2 the_AT jury_NN further_RBR said_VBD in_IN term-end_NN
 presentments_NNS that_CS the_AT City_NN Executive_JJ Committee_NN ,_,
 which_WDT had_HVD over-all_JJ charge_NN of_IN the_AT election_NN ,_,
 deserves_VBZ the_AT praise_NN and_CC thanks_NNS of_IN the_AT City_NN
 of_IN Atlanta_NP for_IN the_AT manner_NN in_IN which_WDT the_AT
 election_NN was_BEDZ conducted_VBN ._.

SA01:3 the_AT September-October_NP term_NN jury_NN had_HVD been_BEN
 charged_VBN by_IN Fulton_NP Superior_JJ Court_NN Judge_NN Durwood_NP
 Pye_NP to_TO investigate_VB reports_NNS of_IN possible_JJ irregularities_NNS
 in_IN the_AT hard-fought_JJ primary_NN which_WDT was_BEDZ won_VBN
 by_IN Mayor-nominate_NN Ivan_NP Allen_NP Jr._NP ._.

A corpus represents a sample of language and contains a cross-section of text from a variety of genres; or it can be compiled within a particular knowledge domain. The texts that constitute a corpus can be organized into general or specialized corpora. A corpus for either specific or general purposes may contain a few thousand words of text to hundreds of millions of words. Since a corpus represents language production from a defined time and genre, access to such large quantities of authentic data increases opportunities for empirical observation of linguistic phenomena at the word, clause, phrase, and textual levels of analysis (Francis, 1993).

A general corpus is most useful for building up an idea of the global behavior of vocabulary items or phrases. A general corpus can be used for publishing dictionaries and textbooks. For example, all cited examples for interactive dictionary resources on the CD-ROM, the *Collins Cobuild English Grammar* (Sinclair, 1990), were drawn from the Birmingham Corpus materials. This demonstrates that the reference publications depend on the provision of frequency information. Unlike this corpus for dictionaries, a specialized corpus consists of language selected for specific purposes. A specialized corpus is particularly useful for understanding how language is used in a particular register of the language (Tribble, 1990). For instance, a corpus that consists of news articles on a certain world event such as a war can help journalists understand a particular group's perspectives at a particular time. A corpus that consists of creative writing allows for an analysis of images and metaphors that were applied by specific writers or facilitates comparative studies on writers. Similarly, a corpus that consists of business letters can provide an insight into the grammar and types of wording that are appropriate to a variety of commercial transactions.

A learner corpus is another example of specialized corpora. A learner corpus has a variety of subcorpora, which represent different types of learners. In the study by Karin Aijmer (2002), 200,000 words constitute a corpus consisting of four subcorpora, each of which was taken from argumentative texts of university students with an advanced level

of proficiency in English: the LOCNESS (NS corpus), the ICLE Swedish, the ICLE French, and ICLE German. Aijmer encourages researchers to interpret quantitative results cautiously because native speakers and nonnative speakers express modality differently, depending on the topics and text type, such as literary versus argumentative.

Additionally, subcorpora play an important role as a resource tool. These can aid a language teacher in finding items that are typically over used, underused, or misused by nonnative speakers (Leech, 1997). Despite the advantages of using corpora for ESL education, especially for English for Academic Purposes (EAP), they have been relatively underused in language teaching and learning (Thompson, 2002). Relatively smaller or midsize corpora can be used to meet the interests and demands in academic and professional settings (Curado, 2002; Scott, 2000a).

Transfer problems or the general strategies that L2 learners use may also lead to an over- or underrepresentation of a specific structure or word in a NNS corpus. Additionally, language-specific cultural norms, differing rhetorical traditions, or ways of teaching composition can also be reflected in a nonnative speaker (NNS) corpus. For example, some cultural groups value directness and certainty in academic writing, whereas others prefer a more indirect style. In response to this, the native speakers' (NS) essays were used as a backdrop with which characteristics of learners' use of modal expressions could be compared in terms of underuse and overuse of modal forms and meanings.

In the study by Aijmer (2002), it was found that modality in advanced learner writing was generally overused in all formal categories of modality, and underuse of modality was only found to be at the functional level. The corpora of NNS usually contained language that is more speech-like, when compared to the native English writing samples. Implications show that future work on learner writing using modality should control the topic more, as it has been a key factor in modal use. Additionally,

more learner groups at varying proficiency levels should also be investigated in order to have a clearer picture of the use of modality, specifically, in learner corpora.

Such varied levels of analysis through observation and evaluation of the facts of language shown in authentic language production can be utilized in second language learning and teaching (Murphy, 1996). For example, second language learners who are studying to become lawyers where their second language is spoken will need to learn domain-specific vocabulary and expressions. They could use corpora that contain domain-specific terminology needed for their subject area. Similarly, in the case of English language learning for specific purposes, the corpus should be compiled with greater specificity. This corpus is likely to have a smaller number of texts adequate for ESP applications (Flowerdew, 1996). A corpus should contain appropriate examples of language production of the genre which learners are interested in. This corpus can then provide learners with important resources such as essential information about the lexicon, grammar, patterning and organization for the genre (Tribble, 2001). Leech (1997) refers to this as a "sub-language." On the other hand, if a corpus is compiled to be used for English as a second language (ESL) classes for general purposes such as survival English, it is important that the corpus consist of a wide variety of texts.

Second language learners can use corpora through a concordancing program because corpora provide authentic and naturalistic data. Through observing and evaluating the data, students can develop language awareness (Carter, 1993; Tribble & Johns, 1990). As a result, access to a corpus and concordancing program helps the learners to have a wider understanding of the language learning process. It also has a potential application in the language classroom by employing general principles and methods of corpus-based language analysis (Murphy, 1996). For this reason, corpora and concordancing programs have been used by second language learners and teachers in classroom exercises. These exercises include building vocabulary and exploring grammatical and discourse features of texts (C. Kennedy & Miceli, 2001). For example,

Aston (1997a) suggests an exercise using a corpus that consists of several texts on the same topic. By using this specialized corpus, the learners can retrieve multiple texts that contain recurrent patterns for analysis. The texts retrieved from such a corpus can serve as a source for discourse analysis. If learners understand the context, they can also be asked to produce texts that are similar to those retrieved from the corpus. In this case, a corpus can provide examples with particular collocations in particular situations.

Minugh (1997) encourages the use of newspaper CD ROMs as corpora, pointing out that contemporary newspaper CD ROMs are an excellent resource for new words or phrases. When using these corpora, it is possible to find many examples even if the words or phrases may not appear to be popular. Secondly, newspaper CD ROMs make it possible to see the extent to which a word or phrase goes from one dialect to another. Many EFL students need this variety—particularly if they are taught by UK-oriented teachers and use US produced class materials such as books and videos.

An example will indicate how a newspaper CD ROM can be used in the classroom for studying grammatical constructions and specifically, collocations. A minimal search in *the New York Times* for *high time* from January to March 1992 provided 13 examples. A few of those are as follows:

- high time + to-infinitive: 5 examples
EX: “Whatever the fate of this year’s nominee, it is in the party’s best interest to change the system. It’s high time for the Democratic Party to rediscover Democracy.” (3 March)
- high time + subjunctive infinitive: 1 example
EX: “Consequently it is high time that ethnic Russians, those who suffered most under Communism, be given preferential treatment and that a strictly Russian ethnic quota be restored to remedy seven decades of injustice.” (20 February).

It is also possible to examine style and register with newspaper CD ROMs. Various text types, such as economic news, editorials, sports, and the like, can be used to compare grammatical and lexical features of these subgenres. The expression *dark horse*

could be tested and found to be a political metaphor or racing term. However, newspaper corpora may have a somewhat narrower range of registers than other corpora, such as those consisting of texts from fiction.

Corpus linguistics can be used as a study tool in conjunction with a dictionary to find authentic examples in learning a foreign language. It can also be used as a reference tool, especially in composition, to find the appropriate use of a word in a sentence and surrounding words. Using a corpus for second language learning and teaching will be discussed in further detail in Section 4.

2.2.2. Concordancing

A concordance is a list of all words found in a specific text or set of texts. It shows the larger context where each word is found. Concordancing is a way to access a corpus of texts in order to show how a word or expression was used in the given context (Flowerdew, 1996). Computer-based concordancing programs perform this analysis automatically. Such programs as *Longman Mini-Concordancer* (Chandler, 1989), *MicroConcord* (Scott & Johns, 1993), *WordSmith* (Scott, 2000b) and *MonoConc Pro* (Barlow, 2002) provide an option for users to retrieve contexts in which the search words or phrases occur. These contexts can be displayed in various ways. The advantages of using computers for corpus linguistic investigations include automatic searching, sorting and scoring. For example, *MonoConc Pro 2.2* (Barlow, 2002) looks for the search word in the corpus and displays the results as it is performing the search. It is also capable of counting word and collocation frequencies.

Figure 2.3. Concordancing program: *MonoConc Pro 2.2* (Barlow, 2002) and *English Gigaword* (Graff, 2003)

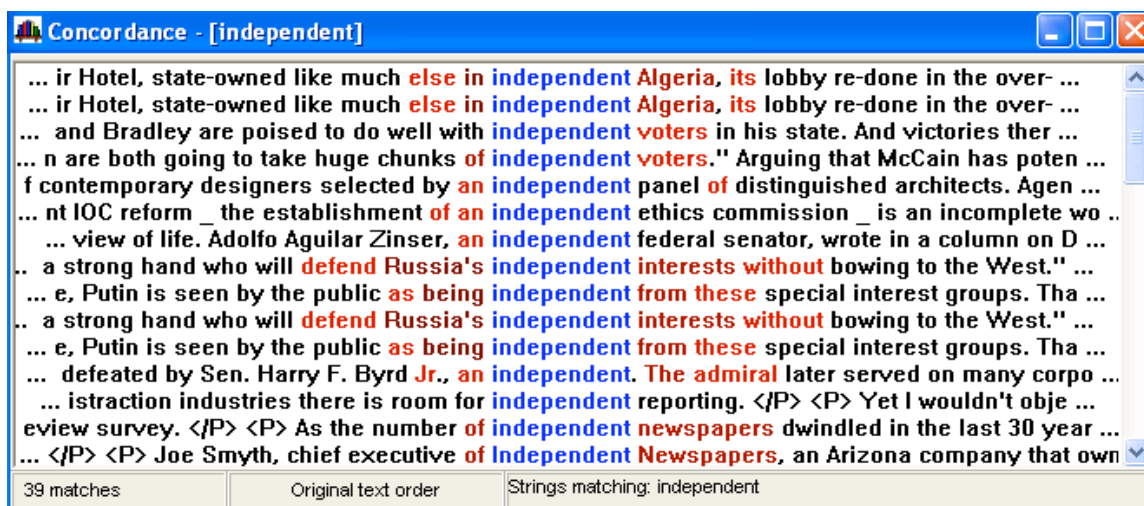


Figure 2.4. Concordancing program search results: *MonoConc Pro 2.2* (Barlow, 2002) and *English Gigaword* (Graff, 2003)

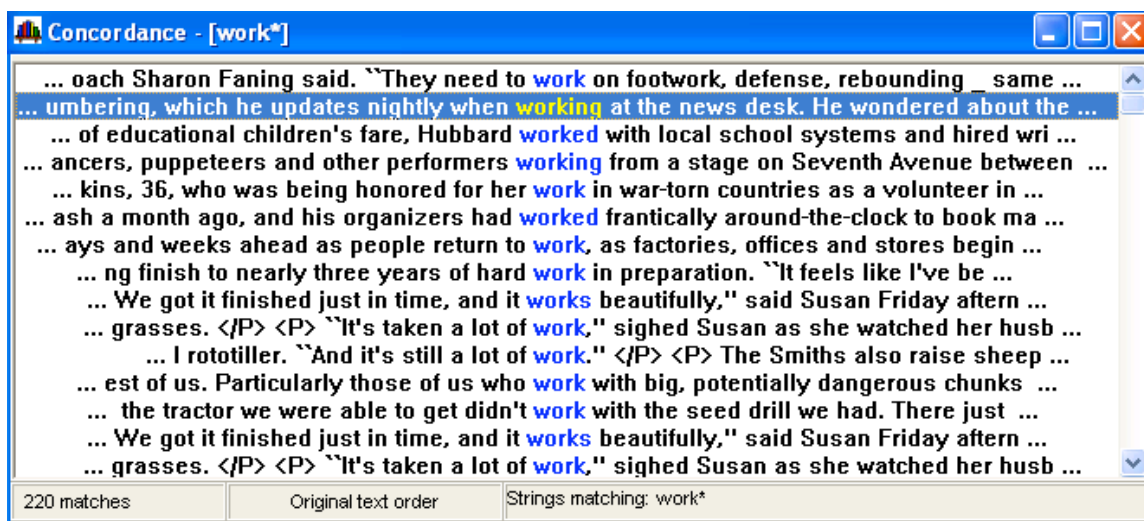


Figure 2.5. Concordancing program collocation frequency: *MonoConc Pro 2.2* (Barlow, 2002) and *English Gigaword* (Graff, 2003)

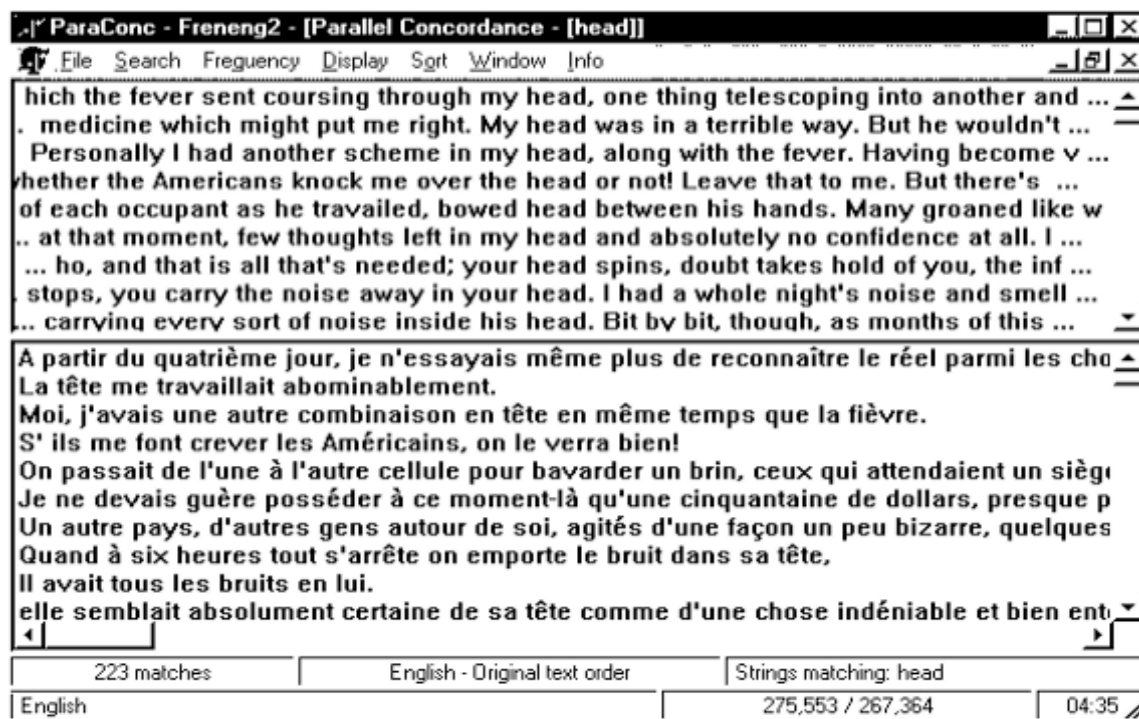
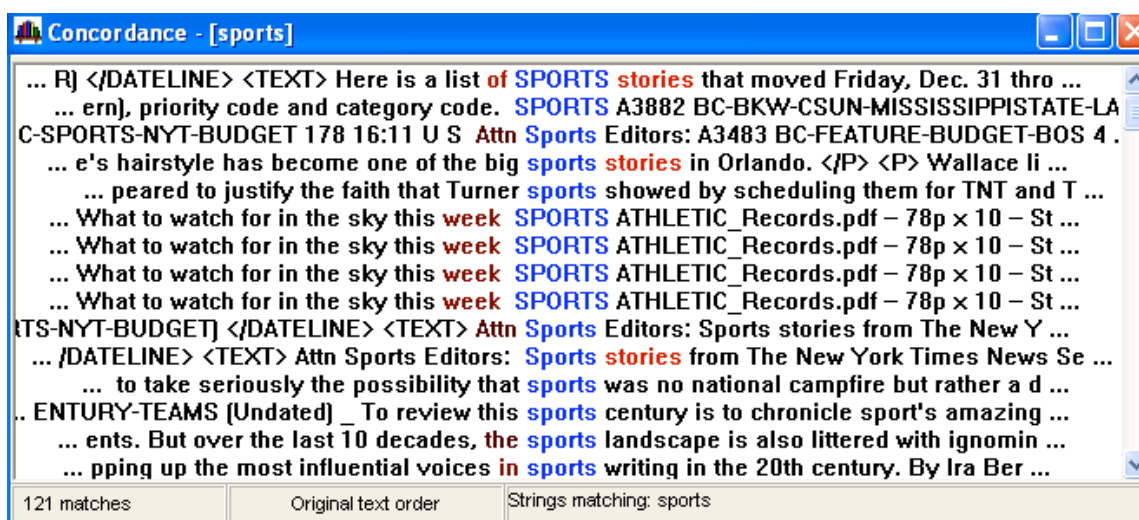
1-Left	1-Right
26 to	34 in
23 of	30 for
22 2000	22 and
17 the	19 on
17 who	18 to
13 a	18 with
12 find	16 as
10 are	15 at
10 not	13 ''
9 and	13 of
9 had	10 who
8 their	9 by
7 it	8 out
6 continued	8 but
	7 ...

In order for these procedures to be carried out, a corpus has to be in plain text format. In other words, a corpus has to be readable by a concordancing program, which will find each occurrence of any requested word and will also place it in context. The amount of context varies according to the settings the user chooses or the software that is used. Usually, only one line of text is provided, which may or may not be a complete sentence (see Figure 2.3). Concordancing programs can search not only for a single word, but also for all occurrences of a word with a specific stem. This can be done using special characters such as the asterisk (*) for partial words or the 'at' symbol (@) for any number of words occurring between two search words. For example, a search for the word itself only produces results with "work" only. Searching with "work*" will produce results with "worker," "working," and "worked," etc. in context (see Figure 2.4). Users can also type in groups of words or phrases on the computer screen. The concordancing program then displays the most typical patterns of the given words or phrases. A concordancing program is able to provide additional information such as word frequency lists, collocation frequency, an alphabetical list of all words in a corpus, and the number

of times a word occurs (Murphy, 1996; Wichmann, 1995). Figure 2.5 shows the collocation frequency list of “work*.” It indicates that “to” occurs 26 times right before the search word (“work*”), and it is the most frequently occurring word in this position. The figure also indicates that “in” occurs 34 times right after the search word, and it is the most frequent word in this position.

Some concordancing programs such as *ParaConc* (Barlow, 2001) can display search results in multiple languages (see Figure 2.6). One purpose of such programs is to permit investigations of translated texts. *ParaConc* (Barlow, 2001) allows loading of any language pairs. For example, this program can be used for English-Chinese or French-Italian texts. It accepts as many as four different languages. This means that the program can display search results in four different languages. When a user clicks on a line in the results window, the corresponding segment in different languages will be highlighted. The following figure shows the search results of “head” from an English corpus and corresponding text segments from a French corpus in the lower window.

Before going into detail about the use of concordancing programs, it is necessary to explain how a corpus can be utilized by a concordancing program. The Key Word In Context (KWIC) format is the most popular form of output. For example, a search word is shown in the middle of the line, surrounded by authentic contexts. The scope of the context words can be specified by the user (see Figure 2.7). Usually just a single line of text is provided, which may be a sentence fragment. The KWIC format enhances understanding of the key word by providing reliable and realistic contexts in which the word is used (Murphy, 1996). In some concordancing programs, a specific word order can be used as a search term.

Figure 2.6. *ParaConc* (Search word “head”) (Barlow, 2001)Figure 2.7. Concordancing program search results in KWIC format: *MonoConc Pro 2.2* (Barlow, 2002) and *English Gigaword* (Graff, 2003)

Kennedy and Miceli (2001) asked students to use a corpus while they revised their own written work. The class was presented with anonymous sample sentences from the previous week's writing at regular intervals. They worked with this assignment in order to practice ways of using the corpus for correcting their writing. The goal of this research was to give students a corpus to use as a primary reference tool while writing. A corpus appropriate for their proficiency level and tasks was given in order to provide examples of personal writing on general topics.

Dodd (1997) studied the use of a corpus in a language classroom. The students were given a new raw corpus, which they used to compare with reference works. Dodd's conclusion was that a computer-supported investigation is a powerful but simple tool for language learning. This supports Leech's opinion (1997) that a computer promotes a learner-centered approach with an open-ended supply of language data that encourages discovery learning.

In addition to serving pedagogical purposes, concordancing programs have played an important role for linguistic and literary researchers (St. John, 2001). For example, Mintz, Newport, and Bever (2002) investigated linguistic input directed at young children under two and a half years old. They wanted to find out whether this input contains adequate information for the acquisition of grammatical categories of noun and verb. Input corpora selected for the study was from the CHILDES database (Mac Whinney, 2000). The corpora selected from this database contain a significant number of utterances commonly directed at children under two and a half years of age. Two hundred of the most frequent words were used. Low-frequency words were not used because they include very few contexts applicable to this particular age group. The results demonstrated that the information in the input could help in constructing grammatical categories of nouns and verbs. Additionally, they discovered what type of information must be available for learners to categorize nouns and verbs. This study supports nativist claims for the existence of innate lexical categories, which, according to

the theory, all humans possess, which makes it possible to categorize lexicon at even a very early age.

Concordancing has various applications. Lexicography and dictionary making were the first applications of concordancing (Flowerdew, 1996). The *Collins Cobuild Dictionary* was a result of authentic concordancing examples. This corpus has also been utilized for literacy and linguistic research as well as stylistics. As stated previously, these applications could potentially contain millions of words. These tools challenge fundamental linguistic descriptions because they allow for explanations based on evidence instead of intuition (Sinclair, 1986). Observation of language data seems to be the most reliable source of evidence for certain types of language phenomena, such as frequency, while Chomsky may suggest that this type of quantitative data is meaningless (McEnery & Wilson, 2001).

In addition to the aforementioned benefits, concordancing is extremely efficient and has great potential for innovation. Concordancing was characterized by Stevens (1990) as "economical in terms of time" to carry out text manipulation because it only requires a program and a collection of texts. For example, concordancing can be used as help for computerized cloze exercises so that students can find more about the nature of the word in the gap (Stevens, 1995).

2.3. Dictionary use

In this section, L2 learners and learners' usage of dictionaries (monolingual, bilingual, and bilingualized), will be discussed. The three parts include general use of a dictionary and how L2 learners benefit, problems with dictionary use, and how corpus usage complements dictionary use.

2.3.1. Dictionary use and benefits

Nation (2001) defines dictionary use as the "deliberate, explicit study of words." Traditionally, paper-based dictionaries are one of the main sources of information about

words that students can use. Paikeday (1993) considers examples in a dictionary even more important than definitions. These examples become ideas. When one uses a dictionary to turn ideas into language, it is called encoding. Some professionals suggest using a monolingual dictionary, while others support bilingual dictionary use, and still others recommend using a combination of both. Bilingualized dictionaries are recommended by other language professionals. Bilingualized dictionaries provide information about word entries in the learner's target language and their translations in the learner's L1.

In order to make a better judgment about dictionary use, it is helpful to look at other researchers' findings. The aim of Knight's study (1994) was to find out whether there is a significant difference in students' vocabulary test scores when words appear in context versus words appearing without context. Knight was interested in finding the difference in vocabulary scores between high-verbal ability language students and low-verbal ability students, as well as the difference between vocabulary learning scores of students who used a dictionary and those who did not.

The participants consisted of 112 intermediate college-level students of Spanish. The students were asked to give definitions in an open-ended test and choose the correct definitions from a multiple-choice test. One group of students used a computerized dictionary on both of these tests, while the control group did not. This dictionary used was programmed to count the number of times the students looked up each word and the amount of time spent on each word. The students who had access to the dictionary pressed a key on the computer and typed the root form of the word in Spanish. They were then able to see the dictionary definition in the center of the screen if they typed in the correct root form in Spanish. Knight found that vocabulary comprehension and retention was greater for the students who used a dictionary than those who did not. Dictionary use did not appear to interfere with subjects' reading comprehension.

High-verbal ability students performed better overall than low-verbal ability students. The students with low-verbal ability did poorly when asked to guess from the context. High-verbal ability students spent less time looking up words in a dictionary than their low verbal-ability counterparts. The results suggest that the common practice of encouraging students to guess from context should be reconsidered. According to this study, dictionary use should be encouraged for low-verbal ability students unless they are able to derive meaning from context.

Gonzalez (1999) obtained results similar to those of Knight. Gonzales conducted a case study entitled "Building Vocabulary: Dictionary Consultation and the ESL Student." The subjects in this study were ESL high school graduates who were enrolled in an intermediate college reading course. For 12 weeks, the students participated in this study by selecting 36 content-specific articles of at least 180 words and writing a summary or comments about the articles. They were told to choose five words unfamiliar to them from each article that were not necessary for comprehension. For each word, they could guess from the context or refer to a dictionary and then write the definition with a summary. The researcher then categorized each item as noun, verb, or adjective and then compared the lexical definition with text representation or context.

Students were interviewed on the amount of time it took to do the work, problems encountered while working and whether a bilingual or monolingual dictionary had been used. Approximately 79% of the word entries proved to be correct when paired with the text they represented. The words that caused the most problems were generally those that had unfamiliar morphological constructions. The conclusion of this study was that, although dictionary use is difficult and sometimes frustrating, it is necessary. One implication of this study is that beginning ESL students must be taught dictionary skills in order to enhance their vocabulary development. With these skills, they are better able to compensate for their lack of vocabulary by developing strategies that meet their specific needs.

It is important to remember that many dictionaries give not only definitions, but also examples, morphology, and the part of speech. In his 1997 study, Christianson identified what helped students while they were using their dictionaries during the writing process and why. The participants were Japanese EFL students enrolled as freshmen in a Japanese university. They were told to underline all the words they looked for in a dictionary for their in-class writing assignments. Their grades were based on the length and extent of their writing development. However, grammatical accuracy was not considered as a criterion for their grade. The students were frequently told to underline words that they had looked up in a dictionary while writing. Because of this, most of the dictionary words were accurately accounted for. Analysis of the words students looked up indicated the students' vocabulary levels. After seeing the words the students looked up, Christianson categorized those dictionary words that caused problems for the students. He then interviewed eight students. Four of these were successful dictionary users and four were not. The results from the interview showed that students who read example sentences were able to apply such information to their writing and made fewer errors. It should be noted that knowledge of an entry's function and use, such as prepositional allocations, noncount, and count, should enhance dictionary use. Christianson also suggests that all dictionary entries should provide useful examples that students can apply to their writing. He concludes that some people still consider dictionaries to be just another learning tool, while others believe that a conscious effort should be undertaken to improve students' dictionary skills and FL writing ability.

Depending on what kind of dictionary is used, language learners can benefit in different ways. Baxter (1980) conducted a study in which he supported the use of a monolingual English learner's dictionary. First, he took into consideration the nature of vocabulary in language and language instruction. He noted that students often felt frustrated because of their lack of words to express themselves. Spoken and written language differs in the time given to refer to a dictionary, thesaurus, etc., for assistance in

a word gap. A writer has sufficient time to search for help with a word, while a speaker must continue with the flow of the conversation if he/she wishes to keep his/her turn.

The group that Baxter chose to study dictionary use consisted of 342 Japanese university students in Japan in 1979. Sixty-two of the participants were majoring in English, while 280 were not. The participants were given a questionnaire on the use of monolingual English dictionaries, bilingual Japanese-English dictionaries, and bilingual English-Japanese dictionaries. (In Japan, the last two dictionaries are usually not combined into one volume.) The majority of the participants, English majors included, chose the English-Japanese dictionaries over the others. Monolingual dictionaries were often criticized for being too difficult to understand. Baxter endorses English learner's dictionaries (dictionaries developed specifically for language learners) as the most appropriate choice for the ESL classroom because students often lack a way to describe a vocabulary term they cannot recall or have not encountered.

Baxter advises that students' feelings should be considered when introducing a monolingual dictionary. Many students have already developed strategies tailored to their bilingual dictionaries, but they also need to be taught new ones in order to make the most of their monolingual dictionaries. Therefore, students will learn that a definition is always possible, even if a single lexical item does not come to mind.

Baxter's study indicated that a monolingual dictionary is more helpful than a bilingual dictionary. He supports the idea that monolingual dictionaries encourage the strategy of paraphrasing, while bilingual dictionaries discourage paraphrasing. Paraphrasing, he asserts, makes up for gaps in productive vocabulary, while bilingual dictionaries encourage expressing a meaning through one single word, which may not encourage the more communicative skill of paraphrasing.

While a monolingual dictionary encourages paraphrasing strategies, a bilingualized dictionary helps comprehension. Laufer and Kimmel (1997) investigated which part of the entry in a bilingualized dictionary EFL learners who were native

speakers of Hebrew read when they consulted the dictionary for unfamiliar words. This dictionary provided definitions and examples in English and translations of entries in Hebrew. The students were divided into three groups — good, average, and unskilled dictionary users. The subjects were administered a multiple-choice test consisting of 15 unfamiliar words that are not taught in high school and university EFL courses. The students were told that there can be more than one answer. Each question had four choices including two answers and two distractors and was designed to find out which part of the entry the students read. If they read only one part of an entry in the dictionary, they would only be able to choose one of the two correct answers. If they read both parts of an entry, they should be able to choose the two correct answers. The results indicated that all learners scored best when they consulted the bilingualized dictionary. The researchers concluded that a bilingualized dictionary is the most useful for such learners.

Positive effects of using a bilingualized dictionary on comprehension and production are also shown in a study by Laufer and Hadar (1997). The aim of their study was to investigate which of the three types of dictionary (monolingual, bilingual, and bilingualized) is most effective for comprehending unknown words and producing them in sentences. The Hebrew participants consisted of 76 high school students who had had 800 hours of EFL instruction, and 46 EFL university students who were not majoring in English. There were 15 test items that the participants were not familiar with. The dictionaries used in this study were *Longman Dictionary of Contemporary English* (monolingual), *The Megiddo Modern Dictionary of English-Hebrew* (bilingual), and the *Oxford Student's Dictionary for Hebrew Speakers* (bilingualized). Participants then took a test consisting of five words on each of three test sheets. On each sheet, the researchers included five words with a monolingual entry, five words with a bilingual entry, and five words with a bilingualized entry. The participants were also asked to produce five sentences with each test item. The results show that the participants scored the highest when they consulted a bilingualized dictionary: Consulting a bilingualized dictionary

resulted in significantly better scores than monolingual and bilingual dictionaries for comprehension, and significantly better scores than a monolingual dictionary for production. This means that the participants comprehended and produced vocabulary better when they were provided with definitions and examples in their L2 and with translations of new words. The results suggest that a good bilingualized dictionary can be appropriate for learners of all levels. It is also suggested that before learners choose the meaning of an entry, they should be taught to make use of all information in the dictionary.

Another setting designed to study the effective role dictionaries play in L2 students' understanding of meaning was designed by Bishop (2000). Students were allowed to use their dictionaries on a French exam. The methodology involved giving a questionnaire to students to complete within 14 days after the end of an exam period. A questionnaire was sent out to students up to two weeks after the exam period. The questionnaire was concerned with dictionary use in a particular exam and not during the entire course. Over half of the students commented that they believed their performance on the test was enhanced by using a dictionary. This qualitative study showed one particular group's opinion of dictionary usage. Students experienced positive psychological effects from using a dictionary and also gave advice from their perspective as students on sensible dictionary use and time management.

Fraser (1999) also supports dictionary use from a psycholinguistic perspective. She encourages inferencing along with dictionary consultation in order to achieve the greatest positive effect on L2 reading and vocabulary learning. Evidence from think-aloud protocols indicates that strategic dictionary use enables cognitive processes necessary for vocabulary acquisition. Students, she suggests, should pay attention to form-meaning connections and practice the words for storage in long-term memory, as well as elaborate on associations with other information.

Supporters of dictionary use in the L2 classroom, in summary, have the following aspects to consider. First, dictionary use may increase vocabulary retention rates. Both low- and high-level learners benefit from using a dictionary that provides not only definitions but also example sentences and collocational information. However, in order to utilize information in a dictionary, students need to be trained in dictionary skills (especially lower-level learners who cannot derive meaning from context). Additionally, advanced learners have an advantage over lower level learners when using a monolingual dictionary. Monolingual dictionaries provide a strategy for paraphrasing, which is necessary for communicative academic language skills. Some language professionals encourage using bilingualized dictionaries rather than monolingual or bilingual dictionaries because they appear to be more helpful when used in tests of comprehension and production.

2.3.2. Concerns about dictionary use

In contrast to these positive results of research studies of dictionary use, negative results have been found in some studies of dictionary use. For example, in the 1970s and 1980s, the prevalent processing model for L2 readers was that dictionaries should only be used as the final option. It was considered an interruption to use a dictionary because it was thought to take the reader out of the text (Carrell *et al.*, 1988; Dubin *et al.*, 1986). For this reason, inferencing was encouraged over dictionary use. Harvey and Yuill (1997) carried out a study in which English as a second language learners used a monolingual dictionary. The aim of the study was to find the reasons why learners consulted this dictionary, and how useful this dictionary was as a source of information for problems the learners have when writing in their second language. The participants were asked to write an essay on a given theme of any length within a limited timeframe. The results showed that the motivation for using the dictionary in order to deal with problems regarding word meanings was higher than that for either grammar or

collocation, which are usually considered essential for encoding. The participants read examples in order to clarify, look for, and locate the meanings even though examples are provided mainly to show syntax, collocation, and context. Reports on dictionary use from the participants showed that many informants did not feel secure enough to use synonyms in their writing even though the dictionary provides synonyms. This was because the dictionary did not provide enough context for the synonym. This results in lack of confidence in writing and leads to more errors when attempting to make small cosmetic changes. The results also show that the participants did not rely on the grammatical coding scheme in the dictionary.

Bejoint (1981) found similar results in his study, the aim of which was to find out what language learners need to know in order to access lexical information and what reference skills they must have in order to retrieve and use that information in a monolingual dictionary. The participants were English language learners at a French university. They were asked questions about their dictionary use. The results show that the participants did not make use of all information about grammar in the dictionary. This means that they did not pay attention to the variety of coding schemes. Some of the questions revealed that they often used their dictionary for reading comprehension and composition. They also revealed that the reason they could not find what they were looking for was unsatisfactory definitions. They felt that some definitions were either too complicated or too vague. Bejoint claims that nonnative speakers, unlike native speakers, sometimes face a problem when trying to understand words completely whose concepts are unknown to them. These words are often culture-specific. They usually have a connotation that is familiar only to native speakers. For better understanding of these words, learners need not only definitions, as in dictionaries for native speakers, but also an encyclopedic type of definition. This study also indicates that in order to make maximum use of a monolingual dictionary, it is necessary to have a relatively high proficiency level.

Hulstijn, Hollander, and Greidanus (1996) designed a study based on lookup behaviors. Marginal glosses and dictionaries were used to test the quality of learning taking place in 78 Dutch students studying French. Three groups of students were tested: those using a bilingual dictionary, those using marginal glosses, and those using nothing to read a text and answer comprehension questions. The students were tested on receptive knowledge of 16 target words. Words that occurred more than once were glossed for one of the groups of students. Students were given three tests (preknowledge, recall and recognition), and a final test in which they had to provide the meanings of 16 words.

First, students read the text and then answered comprehension questions. One group of students was allowed to use a French-Dutch dictionary. All students were told they would be tested on their comprehension of the text, but not on vocabulary. The results of this study showed that marginal glosses resulted in better retention scores than dictionary usage. However, the students using the dictionary were more likely to remember the meaning of the words than those using the marginal glosses. The students who had access to a dictionary did not consider it necessary to look up many target words. They only looked up a word if they considered it necessary, were interested in it, or were bothered by the fact that it reoccurred in their text. In conclusion, this study shows that learners have to pay attention to form-meaning relationships. They should pay careful attention to words that are unfamiliar, but may be important. Students should be encouraged to infer and search for meanings, as well as write them down and regularly review them.

Mackintosh (1998) conducted an empirical study of dictionary use in order to gather information on how student translators use dictionaries when translating from a foreign language to their L1. In this study, Mackintosh hoped to gain insight into how dictionaries should specifically be designed, and see how student translators could be better prepared to use dictionaries for translating.

The study took place in the School of Translation and Interpretation (SIT) at the University of Ottawa and involved two tests. The first test was given to eight native English speakers and seven native French speakers. There was a 75-minute time limit for the first test. The participants were asked to translate from their L2 to their L1. They were not pressured to finish the translation although some did. The students were given a choice of four dictionaries: monolingual, bilingual, hybrid general-language, and a specialized paper-based dictionary. Test two was given to 65 native French-speaking students and 41 native English-speaking students. They were given an L1 or L2 definition for each of these nonsense words and asked to give an L1 equivalent for the word.

Implications from test one include teaching dictionary use. Students were encouraged to look up unknown words before rather than during the translation process. Additionally, they were taught to balance thoroughness with time constraints. The second test confirms that L2 students have a good reason for avoiding L2 definitions but should be encouraged to use them as a supplement. Lastly, the idea of creating dictionaries tailored for L2-L1 translation should be explored; electronic dictionaries could meet this need.

Varantola (1998) conducted a similar study, which was a small-scale attempt to analyze the needs that translators have when making an L1-L2 translation of a general text in a specific field. Several hypotheses determined the approach used in her study. Dictionary publishers desire context-free descriptions, whereas dictionary users desire to solve context-dependent problems. Also, translators need reassurance, and they do not like to encounter unfamiliar equivalents. Usually, translators need more than a single lexical item (longer stretches of text). They use the dictionary for nondictionary information in dictionaries because other sources do not contain it. The outcome of a lookup, however, is defined by the skills of the users.

The question Varantola poses in this study is what kind of databases could meet translators' needs along with a dictionary when their needs call for nondictionary type information. The participants of this study were four advanced students of translation. They were asked the following questions: What expression/passage in the original text made you resort to a dictionary? What was your source? Which entry did you choose? Why did you do this look-up?

The data analyzed showed that most searches were in bilingual and monolingual L2 dictionaries. Four look-ups were in a glossary for special fishing terms, one search in a Finnish monolingual dictionary, and another in an encyclopedia. Participants commented that due to the lack of relevant, context-specific examples their bilingual dictionaries were of little help for this translation task. The data showed that among the participants, the one who performed the highest number of searches and look-ups produced the wordiest translation. Individual differences among participants reflected their differing habits of dictionary use and strategies, as well as the different skills in English they possessed.

Varantola concludes that her study was too small for any major conclusions to be drawn. However, she also states that it is still correct to ask whether resources should go towards improvement of dictionaries, improving dictionary skills, or new reference sources which are designed to meet translators' needs (not excluding electronic corpora). The argument, she claims, is that all three efforts are needed. A larger-scale analysis, which would contribute to the teaching of dictionary skills that included details of habits and success rates, would also prove helpful.

The main concerns about using dictionaries are as follows: (1) dictionaries may cause problems in lexical choices when used by novice learners; (2) dictionary formats can cause misinterpretation among all learners; (3) dictionaries may lack a sufficient amount of context to be of any real use to L2 learners; and (4) bilingual dictionaries may

cause certain errors. For these reasons, it is important for L2 learners to be adequately trained to use dictionaries.

Dictionaries may cause problems in lexical choices when used by novice learners. Wichman (1995) mentions that dictionaries seemed to be the most widely used reference tools for foreign language students. She believes that students make incorrect lexical choices that cause problems with their written work. This suggests that either the students lack the necessary dictionary skills or that dictionaries themselves are the root of the problem because they lack information on meaning in context that is crucial to students' L2 needs.

Dictionary formats could cause misinterpretation for all learners. Parry (1991, 1993) investigated the vocabulary acquisition of ESL students who were taking introductory anthropology classes. The intention of the study was to investigate the variety of strategies ESL students used to build vocabulary specifically from their reading over the course of several semesters or years of study at the university. The task of the students was to keep a list of different words from their anthropology textbooks. They were asked to write what they guessed the meanings of the words were. Two subjects who were documented in this study had opposite strategies for learning vocabulary. One subject read and guessed quickly at a few unfamiliar words. This subject checked the meaning in a dictionary, but did not refer back to the context. The second subject read through the readings slowly and stopped at unfamiliar words. This student spent a lot of time on sentences before making a guess and also checked definitions with context. The words that caused the most difficulty for all students included "bridging vocabulary," which is formal prose that expresses relationships or other abstractions (Parry, 1991). Parry concluded that the biggest problem for advanced ESL students is that they will have to learn huge quantities of words they will rarely use again.

Finally, Parry's studies (1991, 1993) arrived at three important findings. First, there is a strong connection between how much people read and the quantity of words

they know. The more varied readings a person undertakes, the more varied their vocabulary will be. Second, people infer meaning from context in different ways. These strategies can influence the way new words are learned. Dictionary use is included among these. And third, the ESL subjects who participated in this particular study tended to misinterpret words regardless of whether they were inferring meaning from context or using the dictionary.

Just as Parry's study demonstrated the necessity for ESL learners to possess strong dictionary skills in order to understand unfamiliar words, it is equally important for learners to have in their repertoire a broad range of semantic categories in order to facilitate the meaning of an unfamiliar word. Effective dictionary consultation requires a user to be lexically and linguistically sophisticated; ESL learners must work very hard in order to achieve this.

Dictionaries may lack a sufficient amount of context to be of any real use to L2 learners. Barlow (1996b) claims that dictionaries do not have a context that is rich enough. Parallel texts, which are a translation of the text in the L2, may be richer, more flexible, and provide patterns that can be searched for. An additional problem that Barlow identifies with dictionaries is that it is difficult to distinguish meanings in the target language that are not distinctions made in their native language. He suggests using the KWIC format in order to effectively examine the context. The structures in these examples can be helpful to determine correct translation choices.

Bilingual dictionaries may cause certain errors. Ard's study (1982) showed the possible consequences of using a bilingual dictionary. This study investigated high-intermediate ESL students' use of bilingual dictionaries while writing. Material used for analysis included a sample composition and an oral protocol that the writer produced while writing. The results indicated that students are unlikely to be able to find acceptable words to use in their compositions due to the nature of bilingual dictionaries. Ard also concluded that a student's native language background also influences the

success rate, since students whose native language is closer to English are more likely to find an acceptable answer.

It is understandable that some of these errors are made because of the nature of bilingual dictionaries. Even when bilingual dictionaries are not used, similar kinds of errors occur. Quite possibly, bilingual dictionaries are used most often when students do not know which word to use. If the goal of an ESL writing assignment is to reduce the quantity of errors, a bilingual dictionary should not be recommended

2.4. Corpora in second language acquisition

To complement the previous discussion of what a corpus is, the design of concordancing programs, and dictionary use, I will discuss the practicalities of corpora with second language teaching and learning. Nowadays, not only are dictionaries a key resource for L2 learners, but computerized corpora are as well. With development of user-friendly concordancing programs and English corpora, attention has been shifted to an approach to using authentic language samples in the classroom (Johns, 2002).

Wichmann (1995), for example, used corpora and a concordancing program in order to teach German. She did this because, in her opinion, dictionaries do not give enough contextual meaning. Her study, however, did not explain the specific activities she had assigned to her students. Aston (1996) also used a corpus (the British National Corpus) in addition to SARA (SGML-Aware Retrieval Application) retrieval software. This software allows the inexperienced user to search quickly for examples of specific words, phrases, and word patterns with the British National Corpus. Results can be sorted and displayed in a variety of formats. The learners reported that they were successful using this concordancing program as compared to "conventional reference instruments" (p. 190). For example, concordancing can provide evidence where the expression is used with particular collocates in a particular situation. In her 2005 study, Yoon had ESL students use a concordancing program as a reference tool for writing in English. She

found that the students used the program as a linguistic resource to check their writing quickly for grammatical mistakes. The program served as a reference tool that provided immediate linguistic help by showing sentences where search words were actually used. The sentences showed how related words were used in different ways in different contexts. The linguistic elements the students checked included prepositional usage, verb and noun connections, subtle differences in meaning, correct word usage and context of use. As a result, the students developed an awareness of lexical use. In the survey given at the end of the academic term, the students indicated that the concordancing program was useful for initial composing and revision of their writing.

Concordancing thus complements a dictionary as a reference tool for writing. Most importantly, corpus use can create an active learning environment. In this section, the discussion focuses on using a corpus for second language acquisition in terms of context, collocation, and active learning. At the end of this section, there is also a discussion on specific techniques of corpus use for second language learning and teaching.

2.4.1. Context

An advantage that a corpus brings to second language acquisition is the context it provides via authentic examples. A corpus-based approach gives authentic examples while providing students opportunities to be explorers (Dodd, 1997). With authentic examples, learners will see the true use of language. Flowerdew (1993) cautions against using fabricated examples due to the possibility that learners may see a "distorted picture of actual use."

Concordances can demonstrate the organization of lexical items and the relationships they have with other words. Such lexical relationships include antonyms, hyponyms, synonyms, and collocations (Carter & McCarthy, 1988). Stevens (1991b) carried out a study to disprove the hypothesis that cloze exercises and concordance-based

exercises can be completed equally well by language learners. The participants of the study included first year male and female Arabic-speaking university students enrolled in an English-for-physics course. Cloze and concordance-based exercises that tested identical vocabulary items were given to alternately seated students during a single scheduled class period. The students were given a set time to complete their exercises. Each of the cloze exercise sentences contained a gap for a missing word. The corpus-based exercises had a set of gapped concordance lines for a single word. The context of the concordance-based exercises was taken from the textbooks. The outcome of the exercises was that the concordance-based exercises resulted in better performance, partially due to the fact that the participants were given more choices from which to pick a correct word to fill in the gap. The format of the concordance-based activity only caused difficulty during a brief familiarization phase.

One conclusion may be that familiar concordance-based exercises can be more easily solved than cloze exercises. However, this is not to say that they are any less challenging than gap-fillers. The concordance-based exercises require students to think beyond discourse fragments and to decide what the fragments have in common. Stevens establishes that students confronted with shortened contexts in concordance output are not hindered and they can discern the word missing from the context. It may even be possible that supplying multiple and unrelated contexts helps learners choose a correct word more than clues in a passage of discourse that all have the same words missing. Stevens also assumes that familiarity played a role in improvement of the concordance-based exercises for both groups compared to those on the gap-filler when the experiment was run the second time.

Although Stevens' particular study showed that readers were able to learn unknown words while reading, other studies contradict these findings. It is difficult to interpret these conflicting findings because researchers do not use the same texts to test the same construct. To test the “guessability” of words, some researchers have created

artificial words or cloze-type blanks instead of real unknown words. Morphemes, therefore, are not part of the testing process. Differing amounts of contextual support are another variation found in studies.

Frantzen (2003) also studied word inferencing, using an authentic short story in Spanish for native Spanish speakers. The text was completely intact and was not altered in any way. The research questions were 1) When a text is authentic and complete, what are the context characteristics that prevent L2 learners from deriving correct word meanings from the contexts? 2) What are the behaviors of L2 learners that prevent them from deriving correct word meanings from contexts when the text they appear in is authentic and complete? 3) What are the patterns and phenomena that appear in L2 learners if inaccurate inferencing arises when learners rely on context?

Results of the study show that students were able to provide correct meanings when the word was used in the context of a story. Previously unknown words, when seen in context, resulted in correct inferences, either because subjects correctly inferred meaning or because they had remembered the meaning after seeing it in context. The findings of the study indicate that assessing context was the primary strategy used by learners. This is also supported by Lee and Wolf (1997). Most of the explanations offered by the learners indicated that context was utilized because at least one aspect of the story was referenced by each learner.

Sean Romano Maddalena (2001) also demonstrated an effect of context on vocabulary learning. He investigated how corpus analysis may be used in a second language classroom to help students understand synonymous words. He decided to address the issue of problematic synonyms when his Japanese high school students were preparing to take the Eiken examination, which is administered nationally in Japan. The participants were six third-year high school students preparing to take the second level of the test. The students chose six synonymic pairs that were grouped under grammatical function and categorized as reoccurring problematic pairs.

The investigation was divided into five stages. The first step was to clarify the grammatical function of each word with the assistance of a Japanese English instructor. Secondly, the students were given a full set of concordances taken from the LOB (Lancaster-Oslo/Bergen) corpus, which contains the target words. During the third stage, the students completed a cloze exercise that used the target words in a second group of concordances found in the LOB corpus. The fourth step involved the students producing sentences they had learned from earlier work. The last step was a group discussion in the L1 (Japanese), in which students formulated their own rules of synonym usage they had discovered by participating in the study.

The study showed that students were able to complete their first activity successfully by using their dictionaries. However, the second task caused the participants difficulty when they were asked to work with unknown words in an unknown context. The third task was inconclusive because it was a cloze exercise with a choice of two answers per item. This task produced results that were difficult to quantify. Discussions with the students after this activity indicated that students were having difficulty with the unfamiliar context of the concordancing program. As Laufer and Shmueli (1997) had previously noted: "Associative context is the most beneficial for retention when it is created by the learner himself." This is what led Maddalena to ask students to create their own sentences that contained keywords. After this, the students were better able to see the differences in synonymic words than they had seen in the cloze exercise. In conclusion, Maddalena claims that concordances may be very important to teachers and researchers, but less important to students. Supervision and guidance are a necessary part of the learning process. Therefore, a combination involving focus on context as Laufer and Shmueli (1997) proposed, must be put into practice in the classroom.

Even though a number of studies show positive evidence of context on second language learning, there are some considerations when using context. A large quantity of unknown words may overwhelm a learner to the point that he or she is not willing or able

to use available cues to his or her advantage (Laufer, 1997; Laufer & Sim, 1985; Liu & Nation, 1985; Paribakht & Wesche, 1999; Sternberg, 1987). Kelly (1990) also questioned the reliability of context clues even with easy texts. Kelly tested the ability of L2 learners to use the context to infer the meaning of unknown words when only one word in the context was unknown. The findings showed that contextual guessing alone does not lead the learner to find the correct meaning.

Nist and Olejnik (1995) researched encountering a word in context and using a dictionary to find its meaning in four dictionaries. They did not find an interaction between context and dictionary definitions. The quality of the definition is what determined the quality of learning. Hulstijn's study (1992) suggests that contextual information be supplemented with multiple-choice glosses. This encourages mental effort by providing a choice of inferences.

Other language professionals have a more negative view of using contextual clues to teach a second language. Schatz and Baldwin (1986) claim that context clues appear to cause confusion in identifying a correct word meaning. In their study, extracts of narratives were used to derive the word meanings; this technique was unsuccessful. Dubin and Olshtain (1993) cited an activity in which participants were asked to fill in the blanks in cloze passages with appropriate words. Some contexts had "low textual support" and therefore did not provide enough information to lead participants to fill in the blanks with the correct words or even synonyms (p. 191). This confirms the negative view that using contextual clues for teaching a second language allows for many interpretations and often yields many more than wanted (Haynes, 1993; James, 1983; Stein, 1993).

If language professionals desire exploration of context via a corpus, it is important to look at what research studies dealing with vocabulary have found. Based on two studies on guessing vocabulary meanings from context, Nation (2001) maintains that at least 95% of the words in a reading should be familiar in order for learners to infer the

meaning of unknown words in context. Context entails parts of speech of words, collocations, references of words, and various forms a word can take. All of these details are enriching to learners' knowledge of words. Anderson and Shiffrin (1980) suggest that both readers and listeners use their world knowledge and analysis of the linguistic context in order to comprehend words and phrases. According to the results of their 1979 study, Gipe and Arnold advise teaching vocabulary by using familiar associations and context, in addition to having students keep a journal of reading behavior. They also found context and definitions to be of more value than synonyms or short definitions. In their 1995 study, Nist and Olejnik discovered that seeing a word in context and then looking at its definition assisted students taking a multiple choice test in choosing an appropriate situation in where the word should be used.

Prince (1996) conducted a study on beginning and advanced learners using contextual cues as a learning tool. Contextual cues are based on the learner's world knowledge or the text. It is one of three categories suggested by Carton (1971), in addition to interlingual (based on L1 knowledge or loan words in the L1) and intralingual (based on knowledge of the L2) cues, which are also parts of learners' knowledge sources. The research involved specifically constructed sentences that learners had to use in order to discover the meaning of a word without an accompanying translation or definition. The students were tested by translating isolated words and filling in blanks within sentences. Both beginning and advanced learners had higher scores when they saw L1 words with their L2 translations than when they were provided with contextual information in sentences. There was, however, a slightly higher score on the sentence completion test in the advanced group who learned from context. This indicates that the advanced learners were able to apply the knowledge gained from the original context to a new situation. On the other hand, the beginners performed better only on the translation task. This is an indication that if contexts are well chosen, they can provide information about grammatical features, typical collocations, situations of usage, and other details.

2.4.2. Collocation

Collocation plays a very important role as a part of a contextual clue. Collocation includes grammar and lexis, which are important in language pedagogy, and can be identified by using methodologies that have been created for corpus analysis (G. Kennedy, 1998). In terms of grammar as an integral part of corpora, McEnery, Wilson, and Baker (1997) examine how corpora are used to meet the needs of grammar at the pretertiary level in the UK. Their conclusion is that a corpus needs to be a part of teaching and that it “presents a means by which grammar teaching may be more effective and more importantly, may be rated more positively by learners” (St. John, 2001). In another study (C. Kennedy & Miceli, 2001), students who worked with a corpus had a better understanding of Italian grammar and higher confidence in correcting their own writing. The students defined the corpus as useful for providing authentic language examples; allowing for exploration of various uses in different contexts; and demonstrating functions of specific words and expressions in different text types.

Moudraia (2001) suggests that the lexical approach to second language teaching is an alternative to grammar-based approaches. Advocates of the lexical approach argue that only small chunks of spoken sentences are completely novel creations and that language consists of meaningful chunks that are combined together to produce a continuous, logical text. Lewis regards chunks as the raw data that language learners use to perceive language patterns that are traditionally called grammar. The term "lexical approach" originated from Lewis (1993), and it is based on the following taxonomy he suggested (1997): words (e.g. house, car), polywords (e.g. inside out, upside down), collocations or word partnerships (e.g. absolutely convinced, community service), institutionalized utterances (e.g. I'll do it, Welcome back), sentence fragments and heads (e.g. As far as I'm concerned . . . , The fact of the matter is that . . .) and text frames (e.g. First of all, Secondly, Lastly).

In order to put the taxonomy of lexical items to work, Nattinger (1980) suggests that teaching should be based on ready-made units which are appropriate for specific situations because this is what language production is based on. For example, using translations of first and second language in chunks as well as taking notice of and reading language collocations and patterns can increase language awareness. Additionally, working with language corpora made by the teacher for use in the classroom or available on the internet such as COBUILD Bank of English ([http:// titania.cobuild.collins.co.uk/](http://titania.cobuild.collins.co.uk/)) or the British National Corpus can be used to research partnerships of words, preposition usage, and style.

Sinclair (1987) and Willis (1990) attempted to design lexical syllabi based on lexical, rather than grammatical principles. Willis (1990) wanted to provide a rationale for lexical-based language teaching and syllabus design. He suggested that instructional methodology should put a special emphasis on using words, meanings, and common phrases and patterns in their most natural environments. Both Sinclair's (1987) and Willis' lexical syllabi (1990) are word-based, but Lewis' lexical syllabus (1993) intentionally recognizes collocations of words, and longer multiword items.

These lexical syllabi challenge current pedagogical practice. First, lexical theory is redefining the role of language description. Second, language learners' needs to perceive and use patterns of lexis and collocation challenge the traditional view of word boundaries. Third, language production is based on the retrieval of larger phrasal units from memory, and not a syntactic rule-governed process. The implementation of the lexical approach involves changes in the teacher's mindset, not radical methodological changes. The activities should be directed toward naturally occurring language and should concentrate on increasing learners' awareness of the lexical nature of language.

2.4.3. Active and discovery-style learning

Dobb (1997) encourages the use of a corpus for the purpose of enhancing active and discovery-style learning in which students use a corpus and create their own grammatical rules. They are able to consult the experts and note the rules on their own. It is very possible that students will be able to discover a fact unknown to them as well as find items that are not always covered adequately in traditional grammar books.

In order for students to get the most out of their research into corpora, Kennedy and Miceli (2001) hold that developing appropriate research habits by observing, using logical reasoning and corpus searching can reduce errors. In this way, corpora can be very enriching to the learning environment by providing a wealth of opportunities for using English and observing the patterns of its use. Additionally, the learners become less dependent on their teachers and their teachers become less concerned with the textbook and more focused on their role as a coach for learners rather than as a language expert (Aston, 1997a).

The Data-Driven Learning (DDL) approach is relatively new in the area of active and discovery-style learning. Proposed by Johns (1991a), this approach emphasizes inductive acquisition as the learners analyze language use and patterns of specific items in a corpus (Johns, 1991b; Tribble & Johns, 1990). Johns observed that the learner becomes the researcher in the DDL approach. The methodology of the DDL approach gives the learner direct access to data so that the learners create their own meanings and uses. They select strategies that enable them to perceive similarities and differences of word items. The actual implementation of the DDL approach in a language classroom changes the dynamics of teacher-student interaction. The teacher directs and coordinates student research, as recommended by Kennedy and Miceli (2001) and Barlow (1996b). Grammar is no longer the largest influence in the lesson plan. Students are in control of their learning and must be given appropriate skills to carry out their language investigation. Wang (2001) noted that parallel concordancing in conjunction with the

DDL approach will enable learners to develop more fully in their knowledge of lexical meaning and use based on their exposure to authentic language.

An example of parallel concordancing and the DDL approach was researched in a case study of students using 10% of the LOB corpus in English to check for language equivalents in German (Wichmann, 1995). In groups and pairs, students were given computer printouts to decide the appropriate German equivalent for the English word. They were allowed to use any reference works available to them. While doing this activity, the students became more aware of the subtle differences in meaning in their own language. Wichmann concludes that using a concordancing program initiates consciousness raising in language learning.

Case studies involving the DDL approach imply that rewriting mainstream teaching materials will provide authentic language for language learning. The DDL approach could be used more actively in the language classroom in the form of concordance-based exercises that transfer meaning of context from the corpus. This is intended to help students develop inductive strategies that will help them outside of the classroom (Johns, 1994).

Johns (1988) outlined the specific reasons concordancing should be used in language learning. The primary reason is that it involves authentic text and active learning. Secondly, the learners are in control of the learning process. Thirdly, research is the metaphor for learning. The learners become more competent as they are given access to linguistic data (Johns, 1991a). McDonough's opinion (1986) coincides with Johns' view that it is more likely that students will learn grammar rules more efficiently if they are given the examples first and then invent their own rules. By using concordancing programs, students are able to research all of the possible contexts for a particular item (morpheme, word, or phrase) and print them out in order to rapidly scan for and compare information.

The DDL approach has great potential for language teaching. First, using a concordancing program can have a large impact on the process of language learning. It stimulates students' inquiries and speculations on the target language. They are able to see patterns and form generalizations of those patterns. Second, the DDL approach affects the role of the teacher, who becomes the director and coordinator of student-initiated research. Many new questions arise, however. How much control can and should students have over their own learning? Is it possible for the new approach to be incorporated into more traditional and familiar methods? How much assistance can and should be given to students in order to develop acceptable research methods? The third major effect the DDL approach has had is a re-evaluation of how language is learned and taught. Rutherford (1987) contends that the DDL approach makes a new style of "grammatical consciousness-raising" possible.

Johns (1991a) warns that this language-teaching approach may not work well with all students. Only those who are intelligent, sophisticated, and well motivated are good candidates for this type of methodology. Johns (1988) collected data that included questions such as "What is the difference between 'therefore' and 'hence'?" and "Why aren't all shoulds real shoulds?" If a teacher is not prepared to answer such detailed questions, he/she becomes the research organizer and the students often prove themselves to be better linguistic researchers than the teacher.

The implications of the DDL approach indicate that teaching materials need to be revised and come closer to mirroring authentic language use. If teachers and students are not completely comfortable with the DDL approach, it can still be incorporated into a more teacher-centered setting. There is a wide variety of concordance-based exercise types that are easily transferable to suit different learning styles and needs. This leads learners to develop inductive strategies that also help them become better language learners outside of the classroom (Johns, 1991b).

The previous sections have concentrated on context, collocation, and active or discovery learning. The following sections will involve strategies and techniques for using corpora in language learning and teaching.

2.4.4. Strategies and techniques of corpus use in second language learning and teaching

Because concordancing programs allow for the retrieval of multiple contexts, they can complement dictionary use during writing. For example, if an expression with a specific collocate or situation is needed, using both kinds of reference tools can increase the probability of a successful selection (Aston, 1997a).

As students search via a concordancing program, they should learn to look for regularities in multiple reoccurrences. Most learners are accustomed to searching for one example, so this research technique may be new to some learners. After viewing and reviewing a variety of texts and contexts, learners are able to infer declarative schemata (meaning). They need to be taught how regularities of the target language can be found in corpora and the relevance of searching for them. Additionally, the regularities students encounter may not correspond to statements they may find in grammar books and dictionaries. Larger and more context-specific schemata may be found in a concordance in order to allow learners to progress beyond the basics. Traditional emphasis in pedagogy concentrates on generalized rules such as using *since* for past time. However, learners need to practice identifying collocation and colligation (grammatical patterns), connotation and discourse structuring. Aston (1996; Aston, 1997a) and Gavioli (1997) emphasize the need for learners to be aware of partial regularities, such as the word *since*. They need to be shown how this information can be targeted through retrieving, selecting and sorting. Through this process, learners become independent in their use of corpora as resources for learning. Dodd (1997) supports this view based on his study of advanced students learning German using an untagged German corpus and comparing their results

with reference works. Dodd's conclusion is that a computer-supported investigation of corpora is a simple but effective tool in language learning.

St. John (2001) further supports the idea of using corpora for autonomous learning because students who take control of their own learning are more likely to react positively to grammatical exceptions they encounter, rather than reacting negatively to the introduction of an exception to a rule. This in turn should motivate them to work harder. St. John also has data on student observation and feedback of concordancing program use. A learner in his study found the concordancing program to be very user-friendly. The student was able to ignore sentences that were too difficult and select the sentences needed for the task. This is what St. John calls the "help yourself" learning strategy.

While concordancing programs enhance autonomous learning, the value of error analysis should not be overlooked. Flowerdew (1996) suggests guided use of the concordancing program. For example, Dodd (1997) says that by observing main and dependent clauses, students can see that the subjunctive form *hätte* is used. Students must discard examples that do not follow the pattern(s) they are searching for.

Another issue involved in language learning strategies is the application of deductive or inductive reasoning. Either may be used in teaching with concordancing programs. A teacher may explain a contrast, such as the difference between the present continuous and present perfect tense, which is the deductive approach. Or conversely, a student may be asked to create a rule to distinguish between two tenses (Flowerdew, 1996). Todd (2001) investigated three areas of language teaching—induction, the use of concordances, and self-correction. Twenty-five postgraduate students who were taking an English language support course at King Mongkut's University of Technology in Thailand made small concordances of lexical items from the Internet. They then accessed the concordances in order to understand patterns and corrected their own errors. The results show that the students were able to induce patterns from the concordances

they created. They were also able to use the patterns to self-correct errors. Todd concluded that learners who are trained to derive valid patterns or self-correct errors can be less dependant on teachers.

Aston (1997a) encourages the need for careful design of corpus-based activities in order to teach learners how to exploit such resources. Variables, such as number of texts and contexts retrieved or ease of access to the corpora, can be manipulated to control activities and difficulty level. Aston makes four suggestions. First, he suggests simplifying the data. While newspaper corpora are an excellent resource for advanced learners, an easier corpus for beginning and lower intermediate students may be created or found by the teacher. It is the teacher's responsibility to verify that the corpora are appropriate and understandable for each level. Secondly, selecting familiar text types and topics may simplify a task because words work with and go with a context of shared knowledge. Most learners, for example, are familiar with short news items that address cross-cultural themes. He also recommends reducing the quantity of data, so that learners will become more familiar with component texts of corpora. However, a small corpus is less likely to contain reoccurrences than a larger one. Multiple examples of features for a particular text type may be presented in highly homogeneous collections to counteract this drawback. Lastly, when the task is simplified, learners may be able to discard some data that are not necessary at the moment if the texts are for personal reading purposes. The type of analysis required will also determine the difficulty level. For example, a collocational pattern search would be "simple," whereas a contextualized interpretation of various texts or contexts would be more "difficult" (Aston, 1997a).

Fuentes and Rokowski (2002) describe how to incorporate a corpus-based approach into an ESP syllabus in a university setting. It is suggested that an exercise should involve an exploration of frequent words in a corpus, since students with access to a corpus can widen and enrich their view of lexical items rather than experiencing only focused interactions in the classroom (Hunston, 2002). Fuentes and Rokowski contend

that using multiple but limited corpora promotes corpus-based thinking. Through corpus-driven classroom activities, students will then be conscious of the need to learn certain words and phrases (Fuentes & Rokowski, 2002).

Corpus linguistics should be the base of second language teaching materials. Jabbour (2001) encourages the production of theme-related texts in order to fit specific groups of second language learners. The example she provides is a group of students who need to read medical research articles. A medical research article (MRA) consists of a title, abstract, and four sections: Introduction, Methods, Results, and Discussion. A reference list and appendices are the sections that follow. Language resources that writers of medical articles use include superscripts, tenses, and collocations. Superscripts perform the function of referring the readers to the reference list of the article. Tenses bring a research argument together by connecting time and action. Verbs such as "report," "require," and "show" are important in an MRA. Collocations that are typically utilized make reference to research results performed by other researchers and usually include words such as "it has been reported that" and "the study shows that," among others. A closer look at the aforementioned patterns indicates that frequent words found in a corpus of an MRA include grammar words, research words, and medical terminology. The interaction of research words with specific grammar words forms a context for the medical terminology. Prepositions such as "of," "in," and "to" are frequent and necessary to process research activities and medical terminology.

As mentioned above, concordancing programs can be used as a source for materials development. Teachers can use them to gather authentic examples of usage to present to students (Flowerdew, 1996). Tribble and Jones (1990) describe how to adapt concordancing programs for general English purposes, while Thurston and Candlin (1998) describe using concordances for academic purposes. Deducing the meaning of words in context is recommended in Tribble and Jones (1990). The learners' proficiency

and learning styles will determine the structure of teaching materials and activities (Möllering, 2001).

Reading strategies suggested by Carrell, Devine, and Eskey (1988) involve an interaction between top-down and bottom-up views of the text. Knowledge of text discourse and knowledge of linguistic elements must therefore, interact. As the learners become more proficient in the language by relating words to meanings, the linguistic elements and growing subject knowledge combine to result in understanding, reading comprehension, and progress in language acquisition.

The methodology used in Jabbour's study indicated that designing teaching materials for medical students involves first identifying the most frequent words in a corpus, which includes grammar words that give way to research words and terminology pertaining to the discipline. Secondly, frames associated with particular words or phrases such as "it + is + adjective + that," as in "It is common that" and "in + noun + with," as in "In patients with." Concordancing programs are able to evaluate real examples and choose the most typical in order to demonstrate common structural patterns of language. Context, therefore, is the key to designing a corpus-driven class or program. Language is not equal to function, but rather contributes to that function. A language item is not an independent entity, but plays a crucial role in a multilayered structure by interacting with other language items that are involved.

2.4.5. Using parallel corpora for second language learning

Pedagogically, a parallel corpus is actually a bilingual concordance. The setup may involve two windows on the screen displaying equivalent textual sections in two languages (see Figure 2.4). Translation equivalences can be compared and researched in the learner's native language or the target language. These corpora can also supply the frequency data of translation equivalence and provide the learner with contrasting grammatical or lexical examples from two languages (Leech, 1997). Barlow (1996b)

claims that corpora and concordancing programs such as *ParaConc* allow students to be investigators in studies relevant to themselves. As language professionals, we can now be confident that language teaching materials are directly related to and current with the language used outside of the classroom. Students can now have their own tools to choose concentrated and authentic portions of language usage.

Zanettin (1994) describes using parallel English and Italian subcorpora via concordancing software. The purpose of the language activities is to develop translation skills. The emphasis in the study is that concordancing programs “can be run by students at any time in a self-access environment, provided that instructional sheets explaining the background for the activity are supplied” (p. 108). Wang (2001) demonstrated the pedagogical application of parallel concordancing in English and Chinese. St. John (2001) described a pilot study designed to determine if a parallel corpus and concordancing program would be appropriate research tools in a beginning-level German program in an unsupervised environment. The beginning German student was asked to find English equivalents unknown to him and create correct grammar rules by himself using the concordancing program and parallel corpus as his only tools. The results show that these resources can be very beneficial for beginners. Barlow (1996b) affirms this idea that the use of parallel texts can at least create an awareness of the target language for beginners. Danielsson and Ridings (1996) also studied the incorporation of parallel corpora (Scandinavian languages and English) into an academic program for training translators.

In this section we have discussed using corpora for the purpose of second language acquisition. Corpora can aid in language learning by providing context that includes collocations by exemplifying authentic language production. These advantages of corpora can create an active learning environment that promotes student-centered learning.

2.5. Concerns about concordancing

This review of the literature indicates that concordancing programs can be beneficial for linguistics, applied linguistics, and literature research, as well as second language teaching and learning. However, there are some things that should be taken into consideration when using a concordancing program. In this section, some concerns and suggestions about using these programs for language teaching and learning will be addressed.

Although corpora use can lead to successful second language teaching and learning, it can also cause problems. For example, Kennedy and Miceli (2001) ran a study in which the students using the corpus became so involved with it that, as a result, they forgot to use their dictionaries. The students would check the incorrect word in the corpus, and decide that it should not be used as an answer. In order to determine their final answer, they relied on their own ideas instead of using their English-Italian dictionaries.

Before implementing a concordancing program in a language classroom, it is helpful to see how it has been used before. However, Cobb (1997) observed that little concrete research had been done into how concordancing programs are used. One reason is that it is difficult to evaluate the success of learning because observation alone is not enough. As of now, there are no established rubrics for checking the effectiveness of corpus use. Moreover, there is no hard evidence that shows what a student is trying to learn from a rich information resource like a corpus. Some students using a corpus may wish to increase their vocabulary, while others may only be checking to confirm their answers. Finally, there are so many variables that it is almost impossible to conduct a controlled experiment. The same student may not use a corpus the same way twice.

Turnbull and Burston (1998) studied the importance of training learners on how to use a concordancing program. Their case study involved two graduate students—one male and one female—who were enrolled in Master's Degree programs. A Total

Correction Approach was taken in this study, which lasted eight months during an academic year in 1996. Students submitted electronic copies of four assignments, which were course-related, for the purpose of correcting their English expression. All lexical, syntactic and stylistic errors from a brief section of each text (approximately 500 words from texts containing 1,500 to 3,000 words) were corrected by the researcher, who used a text revision tool. After the students received their texts, guidelines were given that focused on corrective feedback. The students were first asked to examine carefully the corrections made and to then use the concordancing program in order to focus on language items that were problematic for them. Students underwent a one-and-a-half hour tutorial/demonstration that introduced them to the purpose and functions of a concordancing program.

Students were evaluated in three ways: direct observations, interview questions stemming from observations, and other questions. Data was collected through a tape-recorded questionnaire interview and constructed conference session for each corrected text. Parts of these interviews and conference sessions were also transcribed. Additionally, concordance printouts which were generated (or comments noted) by the students were submitted. After each student's final interview, they were each asked to fill out a written narrative evaluation of their experience.

Insights gained during this study addressed three important questions posed by Johns (1991b) related to introducing students to concordancing work. The first question is how far students who are in a specific educational setting can take charge of their own learning. The experiences varied: one student reported that she had progressed, while the other reported that concordancing data was too confusing. Individual differences in cognitive styles of the two students were also obvious in this study. The student who reported progress adapted quickly to analyzing concordance data. The other student believed that he could best improve his fluency and comprehension through communicative activities. The second question was how much help can and should be

offered to students in order to develop appropriate research methods. The answer to this question was that learning styles must be taken into account as well as students' familiarity with inductive learning strategies. The KWIC format used in this study only displays search results in KWIC format; only limited context can be shown in this format. This results in successful searches only for certain language use that does not require a wide range of context to understand, such as prepositions, conjunctions, participles, and articles. Because of the limited context in which the KWIC format is shown, concordancing investigations for some linguistic items that require context beyond the sentence level may be unproductive. For example, the KWIC format was not suitable for the investigations of content words that did not occur frequently, function words that were used for complex relationships beyond the level displayed in the KWIC format, and expressions for subtle semantic complexities. The two main obstacles presumed in this study had been the limited amount of concordancing training the students had received and the restricted data contexts available to them in the concordancing program. The third question addressed in this study was whether the new approach could be integrated with older and more familiar methods. The study suggested that they could. The integrated system turned out to be more helpful to the student familiar with inductive learning strategies than to the one who was not.

One recommendation to be taken from Johns' study is that an extended corpus should include a student's own corrected texts in addition to a collection of discipline-specific professional texts. In this way, a student will be able to contrast his/her own language use with that of the experts. It is also suggested that the concordancing software used should generate and display data in KWIC format as well as a wider context. Additionally, the results from this study indicate that concordance data alone does not necessarily stimulate inductive learning strategies. The preferred learning styles and experience with inductive learning styles of the students in the study resulted in different outcomes with the concordancing program. This indicates that a gradual

introduction to concordance work and extensive guidance in using concordancing strategies should be provided for the students. Finally, this study concludes that concordancing strategies can be integrated successfully with older and more traditional approaches to learning. However, further research needs to be done to investigate the relationship between using concordancing strategies and language learning outcomes, as well as the relationship between varying degrees of concordance strategy training and learning outcomes.

Ma (1993) also conducted a study of learners using a concordancing program. In order to gather data, he used observations, questionnaires, in-depth interviews, students' diaries, and analysis of students' writing. He concluded that a major requirement for students' success in a concordancing program was training. In this study, the learners' performance of using the concordancing program was poor because they could not utilize the concordancing program to carry out their searches. The learners were not able to make use of the search techniques required to perform searches in a concordancing program as an aid in their writing. This study also showed that learners must be very familiar with the type of information they can discover and techniques for discovering it in order to achieve optimal results in a concordancing program.

Although concordancing programs can be helpful in many language learning and teaching situations, they may not be appropriate in all cases (Flowerdew, 1996). Low- and intermediate-level students may have difficulty reading and understanding context or surrounding words. The readability depends on the sources of texts that a corpus is composed of. If some of these sources do not illustrate a certain usage clearly, the learners may not be able to understand the context clearly. They may not understand the syntactical or semantic part of a search word as a result. Therefore, beginning and intermediate learners may not benefit from working with a raw text or a corpus until they reach a point at which they are no longer concerned about recognizing every language

item in context (Dodd, 1997). This is very common for students who are not able to determine which words are relevant and which are not.

One of the solutions for problems with learner proficiency related to the use of context in corpora is to use a parallel corpus, which is a bilingual equivalent of the text designed specifically for developing translation skills and making learners aware of differences between their L1 and L2. Linguistic researchers have already used this type of corpora (St. John, 2001) successfully. For instance, Wang (2000) developed an English-Chinese corpus and English-Chinese Parallel Concordancer. He (2001) was then interested in trying out parallel concordancing in language teaching. He believed that this type of concordancing encourages learners to be aware of L1 and L2 differences. Learners can compare contexts of the search item in one language with the translation of those in the other language. In so doing, learners can derive meaning from the same contextual information in different languages. He randomly selected 40 examples of a Chinese word meaning *now* and classified them into several groups. Eight Chinese students at a college were asked to give answers for the following questions by analyzing given examples of *now* in parallel texts:

- What underlying pattern can be detected in the parallel texts?
- Why are the English versions of the sentences so much shorter than the Chinese ones?
- What is missing from the English versions of the sentences? Why?

The results show that students found possible structures as well as those not occurring in the examples. They then discovered that some structures occurring in English are not allowed in Chinese and vice versa. This study shows the possibilities of using parallel concordancing for teaching materials to create a Data Driven Learning environment. Wang suggests that language teachers can expect students to use a concordancing program for a particular lexical item and to submit what they found by analyzing the data.

In St. John's pilot study, a beginning student of German used a parallel corpus and a concordancing program in an unsupervised environment. The student was asked to find satisfactory English equivalents to vocabulary that was unknown to him. The first task was to search for words or phrases with a concordancing program and a parallel corpus, which was his only tool. He saw possible responses in the target language parallel to the search language. He had 17 tasks to complete, and if a search word or phrase produced too many hits, he went on to the next task. He only selected sentences he could easily understand on the remaining hits. As a whole, this corpus was too advanced for him, but he was able to strategize and find a level appropriate to his language abilities in the corpora by searching for shorter sentences. The result of the study showed that even beginners can benefit from a concordancing program when translation is appropriate. He used German as his search language in most cases and English was used to help understand the German. However, this student may be atypical, and further research should be carried out that involves learners who use more typical approaches. St. John (2001) suggested that when a larger study is carried out, two groups (typical and atypical students) must be clearly distinguished for more accurate research results. In his study, the subject was a linguistics student who might display different search skills and behaviors from those by students in different disciplines. In other words, he suggested comparing search behaviors by students majoring in other subjects such as science or engineering with those by students majoring in linguistics.

An additional accommodation that can be made for learners at lower proficiency levels is that learners can be paired. In a study conducted by Kennedy and Miceli (2001), students with lower proficiency who were paired with someone of higher proficiency often showed insight in formulating questions, using reasoning for creating a strategy, or paying attention to examples.

In addition to the concerns of learner proficiency level, learner morale should be taken into consideration when using a corpus in a language classroom. St. John (2001)

identified problems in using large corpora for language learners, especially those who are in the beginning and intermediate levels. Students can feel quite unmotivated when frequent words in concordances become too lengthy and meaningless possibly because the amount of information they face is overwhelming. If this is the case, items should be reevaluated for their appropriateness in a concordancing program (Flowerdew, 1996). Aston (1997b) suggested working with a corpus that is small and specialized. In this way, learners can discover the characteristics of a particular area of language use and utilize a small and specialized corpus as a training tool for larger ones.

For example, if learners wish to acquire a certain regional type of English, a general corpus will not be appropriate. A specific corpus of the target area must be used. On the other hand, if the learner desires to learn general English, the language instructor must define what is "general" based on the varieties of genres available in newspapers, periodicals, novels, and other forms of written media. The most typical genre is casual conversation, but, unfortunately, most available corpora are lacking in spoken data, a serious deficiency for learners who want to acquire colloquial spoken English.

In conclusion, the professional assessment of concordancing programs and corpora is that they help organize natural language data so that language learners can see and understand patterns more easily and become competent in the target language more quickly (Stevens, 1995). Tribble (1990) likewise holds that "what the concordancer does is make the invisible visible." In summary, there is an abundance of studies regarding what can be done with corpora in language learning, but little on how learners carry out their investigations (C. Kennedy & Miceli, 2001).

CHAPTER III

METHODOLOGY

The purpose of this research study was to investigate how university students learning English for academic purposes use a corpus as a reference tool in conjunction with online dictionaries. The experimental task was to compose a short text based on a newspaper article.

3.1. Data gathering techniques

Two data-gathering techniques were considered for this study: think-aloud protocols and stimulated recall. The advantage of stimulated recall over think-aloud protocols arises because participants do not have to be trained, nor do they need to be concerned about the influence of simultaneously carrying out a task and thinking aloud. Think-aloud protocol has traditionally been the method used to elicit verbalizations in studying the process of writing for gathering social and cognitive dimensions (Smagorinsky, 1994). Cognitive protocol researchers assume that composing is a problem-solving task, but it could also be viewed as a stream-of-consciousness activity. Another drawback of the think-aloud process is that it may not accurately portray thought processes. It could even be termed subjective, since, while it assists learners in forming a hypothesis, it still does not validate specific cognitive processes. This is why some critics have questioned the reliability of findings using think-aloud protocols (Smagorinsky, 1994).

Stimulated recall is an introspective method used in research to elicit data in order to discover the thought processes involved in carrying out an activity. It is based on educational research and cognitive psychology methodologies (Gass & Mackey, 2000). Primarily, this method stems from the assumption that internal processes may be observed similarly to external real-world situations. It is assumed that people can

verbalize their internal thought processes to some extent. Visual and aural prompts may be used to stimulate the actual recall of events. For example, learners may be given an audio-recording of themselves thinking aloud or a picture they drew or something they wrote during an L2 activity. They are then asked to remember their thoughts and level of motivation during the original event.

Stimulated recall does not entail only one methodology. There are three types that may be used depending on the need. The first is called consecutive recall. For example, after L2 learners finish revisions on an L2 writing assignment, they are interviewed about changes they made. Initial and final written products are used as stimuli. Second, delayed recall may be used in an L2 reading class after participants have read a passage. They may be given comprehension questions asking what difficulties they had when reading the passage and how they solved them. Third, nonrecent recall may be used to research L2 strategies learners have used on a placement test (Gass & Mackey, 2000). For example, students may be divided into higher-scoring learners and lower-scoring learners. Students may respond to the researcher once a week by email and report on how they are learning vocabulary currently, as opposed to during the first half of the year.

Stimulated recall has drawbacks of its own, primarily because it is based on writers' ability to reconstruct their experiences from their own memory (DiPardo, 1994). Participants with a clear memory should be able to simplify, compress information and generalize in order to give an understandable explanation. If some aspects are not remembered, participants may include general information based on their own knowledge of what writers do or they may not include details not related to general writing skills. Another typical problem that participants might have with stimulated recall is that they may not understand key words in questions during the interview process. Additionally, participants may not understand that the purpose of a stimulated recall session is to give

an accurate portrayal of what happened during the writing process, as opposed to giving a "right" answer that demonstrates their competence.

The following study is an example of a successful application of stimulated recall protocol. Sasaki (2004) conducted a study of Japanese students' English writing behaviors over a three-and-a-half-year period of time. She analyzed multiple sources of information including written texts, videotaped writing behaviors, and stimulated recall protocols. Her intention in using stimulated recall was to research the participants' writing processes. The questions that she asked during the session focused on planning, overall organization, moving from one idea to another, translating the generated idea into their L2, and refining the rhetorical aspect of expressions. The participants were asked to finish writing within 30 minutes but were allowed to take longer if they wanted. Sasaki videotaped pencil, hand, eye, and head movements. When they were about to write their very first word, she interrupted them and asked how much planning they had done before starting to write.

Immediately after the participants finished writing in Sasaki's study, the individual stimulated recall sessions were conducted. During the session, they were allowed to explain their actions in either Japanese or English. Her questions focused on finding out what the participants were thinking at the time of the writing session, rather than asking leading questions. She initially asked questions on what their plans were to write at each stage of the composition. She then asked further questions as needed according to the participants' answers. She asked questions every time they stopped writing for longer than three seconds. The types of data collected through the stimulated recall sessions included the time the participants spent before starting to write and during the writing session, and participants' answers to the questions.

According to Gass and Mackey (2000), there are recommendations in four areas that researchers must keep in mind if they are seriously considering implementing stimulated recall as part of their methodology: timing, strength, training, and structure.

This is classified in relation to recall support, stimuli, and procedure. Theoretical predictions and current research have provided a foundation on which to base the following suggestions.

First, because accuracy is crucial to stimulated recall, it is important to do it in a timely manner in order to elicit the most reliable recall. If possible, data should be collected as soon as possible. The more distant in time and memory the original event, the greater the possibility that participants will say what they think the researcher wants them to say. If their memories are unclear, they will invent a likely explanation for their original choices. Garner (1988) discovered that participants who immediately recalled a task had significantly more cognitive events than those who reported two days later. This methodology has an advantage over postactivity interviews because it combines memory with prompts, whereas the interviews alone rely heavily on memory.

Second, the stimulus used during recall should be as strong as possible. It is suggested that two or more sources be used (Gass & Mackey, 2000). An example of this is giving a participant a video to watch immediately following an activity. To enrich the data, a transcript of specific episodes may be included as well. When one of the stimuli is video, the researcher's role is to find out what the participants were thinking at the time of the original event. The researcher may state: "The reason I'm doing this interview is to get a student's perspective. I want to find out what you actually think and do when you write, not what your teacher thinks or expects you to do. Tell me exactly what you did" (DiPardo, 1994). Researchers should write instructions for participants, read them aloud, and possibly record them so that each participant receives standardized instructions. By playing the video, the researcher is able to see participants' actions and listen to their responses, but is not aware of their internal thought processes. An example of elicitation from a researcher could be "Tell me what you were thinking." or "What was on your mind when you were doing that?" When participants indicate that they do not know or do not remember, the researcher should accept the response.

Third, at least minimal training is essential. Participants need to be able to perform the procedure, but cannot be cued on experimental goals or unrelated information. Pilot tests can prevent inaccurate or unhelpful tactics. Easy instructions and a direct model should be sufficient.

Lastly, the structure used depends on the research questions. Researcher interference is lessened if participants are not led (Gass & Mackey, 2000). In order to have reliable data, researchers should not only collect data as soon as the participants finish the task, but should also provide stimuli that help participants recall detailed information (DiPardo, 1994). For this reason, the researcher's questions should ask for concrete examples rather than about writing in general. As a result, the participants are more likely to remember more details. For example, the researcher may ask: "Is there anything else you want to add?" or "Can you explain that a bit more?" The questions or prompts should be clear and based on activities that have already occurred because hypothetical questions can lead to speculations that have little or nothing to do with participants' experiences. Open-ended questions or general prompts should be used such as "What goals, if any, did you have when you began to write?" Asking a question such as "Did you set any of the following goals when you began your paper?" can lead participants to make comments they would not have considered on their own. The questions that will be used to elicit verbal responses should be formulated well in advance. This is due to the need for gathering accurate information without leading the participants to make comments not pertaining to their original thought. Also, for the same reason, specific research intentions should not be obvious to the participants (Lyle, 2003).

When using stimulated recall, it is also important to consider whether the L1 or the L2 should be used for instructions during the stimulated recall (Gass & Mackey, 2000). When recall is done in the L2, the participants' abilities to respond clearly or understand could be inhibited. Therefore, a researcher may have difficulties in reporting

the data because he/she does not know if the response given indicates the subject's true thought processes. Another problem could occur when interpreting what the participants said and attempting to draw conclusions from their responses.

3.2. Pilot study

A pilot study was conducted for two reasons. The first reason was to investigate which of the original tasks under consideration—summarizing and paraphrasing—could create sufficient interactions with the concordancing program. In order to conduct a reliable investigation, it was crucial to gather data that would show a variety of interaction types in quantity and quality. Summarizing and paraphrasing were chosen because these tasks are part of college-level writing and are likely to create interactions with reference tools including online dictionaries and a concordancing program. The second goal of the pilot study was to determine which data-gathering technique, stimulated recall or think-aloud protocols, would best reveal the thinking process of the participants while they were using the reference tools.

3.2.1. Research questions for the pilot study

- (1) How often do the participants consult a corpus when they summarize a newspaper article?
- (2) How often do the participants consult a corpus when they paraphrase a newspaper article?
- (3) Which of the two tasks creates more interactions with a corpus?
- (4) Which of the two data-gathering techniques, stimulated recall or think-aloud protocols, elicits the participants' thought processes better?

3.2.2. Subjects

The participants were four Korean graduate students with high-advanced English proficiency. Three were female students; one was male. In this study, the L1 (Korean)

was used for both subjects and the researcher in order accurately to convey and process information. The data from one participant were excluded because of an unclear voice recording. During the think-aloud session, the volume of this participant's voice was too low to be intelligible and included mumbling that prevented the researcher from retrieving accurate data.

3.2.3. Materials

- Corpus: *English Gigaword*, a corpus published by Linguistic Data Consortium, was used. The participants were given a CD ROM that contained *MonoConc Pro* version 2.2 and text files from *The New York Times* published from 2000 to 2002.
- Dictionary: Subjects were allowed to use online dictionaries: *Yahoo* Korean dictionaries. This site has English-Korean, Korean-English, English-English, and an English-English thesaurus.
- Newspaper articles: Two articles were chosen on current issues: "South Korean Presidential Impeachment (SK) (see Appendix A) and Nuclear Plans in North Korea (NK) (see Appendix B). Both articles shared less than 10% of the same vocabulary.

In order to create equal conditions for both tasks, two newspaper articles of similar length were chosen. Each newspaper article contained 550 words. The following words were not considered to be overlapping because they are commonly occurring words in English.

- Pronouns: I, us, our, he, His, it, that, this, their, they, them, other, another
- Articles: a, an, the
- Prepositions: on, before, about, for, at, of, by, in, with, into, to
- Auxiliaries: will, could
- Conjunctions: but, and, after

- Proper nouns: Korea, Korean

The following words were overlapping words (31 words): *accepting, accused, be, break, come, complete, do, effort, even, force, get, have-has, include, industrialized-industry, just, last, minister, more, named-name, not, nuclear, other, party-parties, power, president, said-say-saying, south, than, time-times, what*. Both of the newspaper articles used contained these words. The intention was to choose two articles sharing a minimum number of the same words. This was done to maximize the use of reference tools for unknown words.

After excluding the overlapping words from the count, text SK contained 394 words, and text NK contained 367 words. The difference in the numbers of words that were actually counted was 27. The number of sentences was the same for both articles. Each article contained 21 sentences. The average number of words per sentence for SK was 26.0, and for NK 26.2. The average number of characters per word for SK was 5.3, and for NK 5.1. Readability for both articles was also similar. The Flesch Reading Ease for SK was 35.4, and for NK 35.0. The Flesch-Kincaid Grade Level for SK was 12.0, and for NK 12.0.

3.2.4. Procedures

There were two meetings that lasted approximately 3.5 hours in total.

- (1) At the first meeting, the subjects (here called Kim, Lee and Park) learned the basic text-searching functions of a concordancing program, *MonoConc Pro* version 2.2, including loading a corpus, frequency of occurrences, single- and multiple-word searches, excluding words from consideration, using wildcards, and displaying contexts.
- (2) The subjects were asked to do the following activities using a concordancing program on a PC: Finding out the word that occurs the most frequently in the

corpus, searching and displaying a word and its surrounding words, searching and displaying multiple words and their surroundings, identifying prepositions used with a certain verb, and figuring out different meanings of synonyms (e.g. 'convince' vs. 'persuade').

- (3) The subjects then were asked to use this program for their own writing purposes for a week in order to get accustomed to using it. They were provided with a CD ROM with *MonoConc Pro* version 2.2 and the corpus.
- (4) A week later, there was another meeting with the subjects for two tasks, summarizing and paraphrasing a different newspaper article for each. Before starting the paraphrasing task, the subjects practiced how to think aloud during writing. They were asked to say aloud what they were thinking while paraphrasing a paragraph-long newspaper article.
- (5) The subjects were allowed to use online dictionaries and a concordancing program during the tasks. There was no time limit for summarizing and paraphrasing. Half of the students did the summarizing task first and then finished with the paraphrasing task. The other half did the paraphrasing task first and then finished with the summarizing task.
- (6) Immediately after the subjects finished their tasks (not including the tasks done with the think-aloud protocol), stimulated recall sessions began. The computer screen was recorded while they interacted with a concordancing program and dictionaries as well as typing text on the screen. The researcher asked questions regarding interactions with dictionaries and the corpus, and the subjects gave answers while watching the recorded screen interactions. The stimulated recall session was recorded to a digital audio file.

3.2.5. Results

Based on the results, paraphrasing a newspaper article was selected as the best task to use. The subjects used the reference tools more often when they were paraphrasing than when they were summarizing. The total use of the corpus was 57 times when paraphrasing, and 41 times when summarizing. The total use of the dictionaries and thesaurus was 51 times when paraphrasing, and 38 times when summarizing. The average of individual participant's use of corpus was 19.0 when paraphrasing, and 13.7 when summarizing. The average of individual subjects' use of the other reference tools was 17.0 when paraphrasing, and 12.7 when summarizing. The total average for individual participants' use of all reference tools was 36.0 when paraphrasing, and 26.3 when summarizing. This shows that the paraphrasing task resulted in an average of 9.7 times more consultation of reference tools than the summarizing task. Two subjects used the corpus more when paraphrasing than when summarizing. Subject Lee used the corpus 13 times when paraphrasing, and 6 times when summarizing. Subject Kim used the corpus 26 times when paraphrasing, and 16 times when summarizing. The results were comparable with the other reference tools. Subject Park used the corpus 1 more time when summarizing than when paraphrasing. This subject used the other reference tool for the same amount of times (see Table 3.1). Overall, the subjects used the corpus more when paraphrasing than they did when summarizing. This was also the case for the dictionaries and thesaurus. In total, they used these reference tools 13 times more when paraphrasing than when summarizing (see Table 3.2 and Figure 3.1).

Stimulated recall was chosen as the appropriate data-gathering technique for the main research study. After reviewing the transcripts from the think-aloud and stimulated recall sessions, it was clear that stimulated recall was a more efficient way to ascertain why the participants were using the corpus, what kind of information they were looking for or reading, and how they used the information for their writing. When thinking aloud,

they did not always say aloud what they were thinking. Even if they did, it was not always clear what information they were looking at and why and how they were using the corpus. For this reason, the data from one participant were not included for analysis.

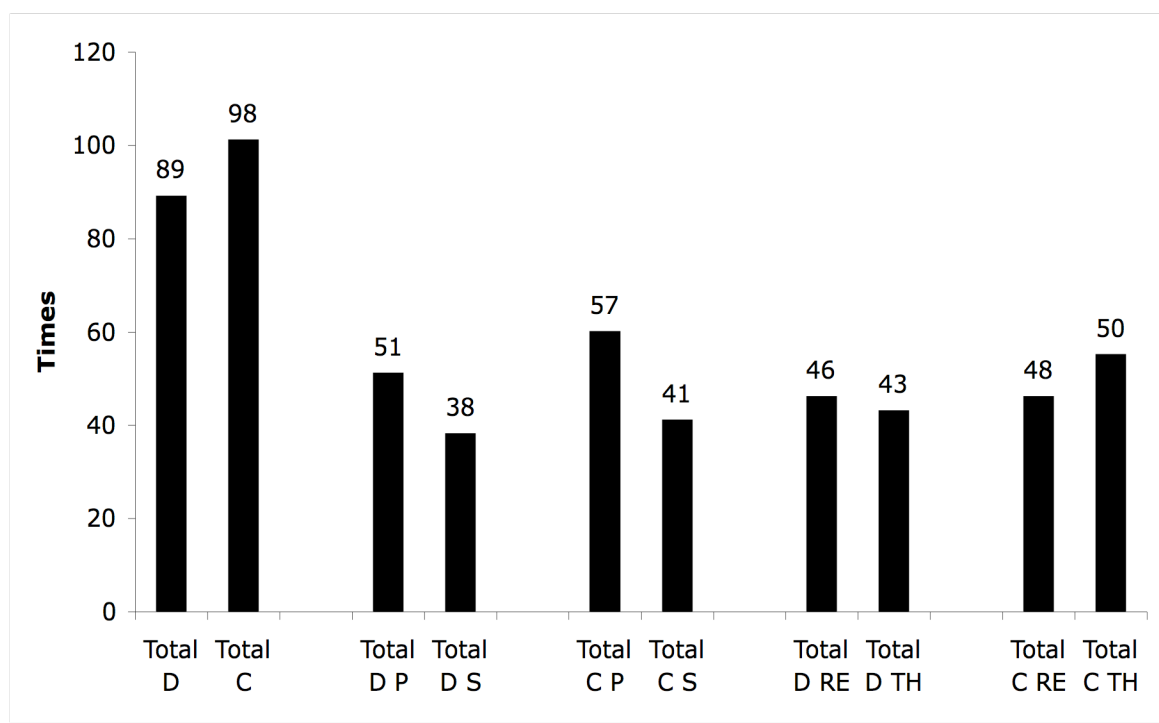
Table 3.1. Individual subjects' use of reference tools when paraphrasing and summarizing

Participant	Task	Reference Tools Used	Number of Use (times)
Lee	Paraphrasing	Corpus	13
		Dictionaries & Thesaurus	25
	Summarizing	Corpus	6
		Dictionaries & Thesaurus	17
Kim	Paraphrasing	Corpus	26
		Dictionaries & Thesaurus	12
	Summarizing	Corpus	16
		Dictionaries & Thesaurus	7
Park	Paraphrasing	Corpus	18
		Dictionaries & Thesaurus	14
	Summarizing	Corpus	19
		Dictionaries & Thesaurus	14

Table 3.2. Use of reference tools for different tasks

Tools	Task	Number of Use (times)
Corpus	Both	98
Dictionaries & Thesaurus	Both	89
Corpus	Paraphrasing	57
	Summarizing	41
Dictionaries & Thesaurus	Paraphrasing	51
	Summarizing	38

Figure 3.1. Use of corpus and other reference tools



D: Dictionary, C: Corpus, P: Paraphrasing, S: Summarizing, RE: Recall Protocol, TH: Think-aloud

In contrast, when stimulated recall was used, accurate information pertinent to the study was gathered more readily by means of carefully crafted questions that allowed the participants to clearly express their thought processes during the writing session.

Additionally, this technique allowed the participants to concentrate on the tasks at hand during the writing session, which included some or all of the following: reading, paraphrasing, translating, using dictionaries and a corpus, and writing. The think-aloud protocol could have hindered subjects' abilities to carry out these tasks while they were expected to give a clear verbal account of the process.

In the end, there was not a large discrepancy in the use of reference tools for the two methods (see Table 3.3). The corpus was used 2 fewer times when stimulated recall was used than when the think-aloud protocol was used. The results were the opposite in

the case of the dictionaries and thesaurus. The participants used these tools three more times when stimulated recall was used than with the think-aloud protocol.

Table 3.3. Use of reference tools for different methods

Tools	Method	Number of Use (times)
Corpus	Both	98
Dictionaries & Thesaurus	Both	89
Corpus	Stimulated Recall	48
	Think-Aloud	50
Dictionaries & Thesaurus	Stimulated Recall	46
	Think-Aloud	43

3.3. Main study

This section describes the subjects' characteristics in terms of computer and dictionary use for writing in English. It also lays out the materials and procedures that were used in the study.

3.3.1. Subjects

The subjects of this study were 10 Korean graduate students at the University of Iowa majoring in different academic fields. Their English proficiency levels were high enough to be enrolled full time in an American university. Seven students were recruited from the Writing Center at the University of Iowa, which offers tutoring to any students wishing to sign up. The other three learned about the research through word of mouth and volunteered to participate. Students using this program were interested in developing academic and personal writing skills. The tutors at the Writing Center worked with the students to help them improve their researching, organizing, drafting, editing, and revising skills. Since the subjects were recruited from the Writing Center, it could be

assumed that they were all motivated to improve their writing skills. The subjects have been in the United States from one to five years. For questions regarding their computer use, all of them use computers several times or more a day for personal purposes and school work. When they use computers for personal purposes, the use of Korean ranges from almost never to about 75% of the time. All of them use dictionaries for writing in English. Some of them only use English-English or English-Korean dictionaries. Others use a combination of both. All of the subjects use paper-based dictionaries. Several of them also use online dictionaries. On the other hand, they rarely refer to the thesaurus when writing in English. Nine of the subjects had never heard of a corpus and have not used the Internet as a reference tool for composing or revising a paper or any other types of writing. One subject had heard of a corpus but had never used one. This subject had used a web search engine as a reference tool.

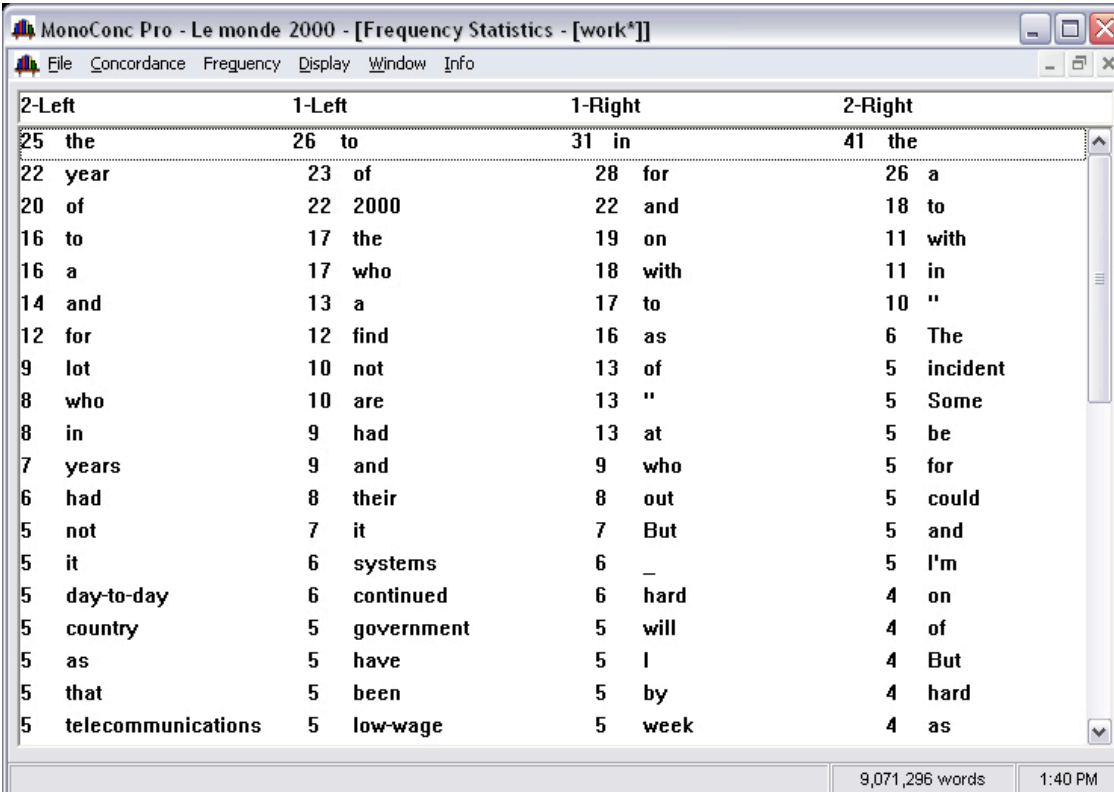
3.3.2. Materials

The corpus chosen for the study was *English Gigaword* (Graff, 2003). It is published by Linguistic Data Consortium. The data of this corpus consists of English news articles from *Newswire*, which includes *Agence France Presse English Service (AFE)*, *Associated Press Worldstream English Service (APW)*, *The New York Times Newswire Service (NYT)*, and the *Xinhua News Agency English Service (XIE)*. The entire content of the *Agence France Presse* is from the years 1995-2001. The content of *XIE* and segments of the *NYT* and *APW* are from February 2001, onward. For the purpose of this study, only the corpus from *The New York Times* was used because the text the subjects were asked to paraphrase was a news article from a North American newspaper.

The concordancing program that participants used to analyze the corpus was *MonoConc Pro* version 2.2, published by Athelstan. This concordancing program displays the results in KWIC (key word in context) format and can count and display word and collocation frequencies. Figure 3.2 shows the output of a search on the word

“work*” and lists the frequency of surrounding words (two words on the left and two words on the right). The users had the option of choosing the number of surrounding words for calculation of their collocation frequency.

Figure 3.2. Collocation frequency statistics (search word: work*)



2-Left	1-Left	1-Right	2-Right
25 the	26 to	31 in	41 the
22 year	23 of	28 for	26 a
20 of	22 2000	22 and	18 to
16 to	17 the	19 on	11 with
16 a	17 who	18 with	11 in
14 and	13 a	17 to	10 "
12 for	12 find	16 as	6 The
9 lot	10 not	13 of	5 incident
8 who	10 are	13 "	5 Some
8 in	9 had	13 at	5 be
7 years	9 and	9 who	5 for
6 had	8 their	8 out	5 could
5 not	7 it	7 But	5 and
5 it	6 systems	6 _	5 I'm
5 day-to-day	6 continued	6 hard	4 on
5 country	5 government	5 will	4 of
5 as	5 have	5 I	4 But
5 that	5 been	5 by	4 hard
5 telecommunications	5 low-wage	5 week	4 as

This concordancing program can load multiple files of text to constitute a corpus. New files can be added or removed at any time. Corpora in any language may be used. Search terms may include wild-card characters such as an asterisk (*) for any number of characters attached to a search word, or others such as percent (%) for zero or one character, question mark (?) for exactly one character, and the at sign (@) for a range of words. Users can save the concordance output to a file and print it out. They can also search for a single word or an entire phrase by means of wild-card characters.

The dictionaries the students were allowed to use in the study were *Yahoo* Online Korean dictionaries (<http://kr.alldic.yahoo.com>). On the homepage of this website, the user can place a search word and select among six options (see Figure 3.3). The first option is English-Korean and Korean-English. If an English search word is used, a Korean translation and a definition with English examples and Korean translations are provided. The second option is English-English. A search word in English is typed in, resulting in a list of related words. The user chooses an option from the list. The next page displays the search word along with the option of an audio clip that provides the pronunciation of the word, along with a phonetic spelling of the word, its part of speech, multiple definitions, some examples of how the word is used in a complete sentence, inflected forms of the word, and its etymology. A thesaurus is included as an option within the English-English dictionary.

The newspaper article chosen for paraphrasing, "North Korea Rejects U.S. Demand to Scrap Its Nuclear Programs," was taken from *The New York Times* and was published on March 28, 2004, and was 550 words in length. This topic was chosen because it was familiar to many Koreans. This particular article was chosen because the corpus in the concordancing program also came from *The New York Times*. Because this article was published in 2004, it was not included in the corpus. Originally in the pilot study, two newspaper articles on well-known issues in South Korea were chosen to be paraphrased. Based on the results, it was established that the subjects used *MonoConc* more when paraphrasing the article about the North Korean nuclear program than the article about the impeachment of the South Korean president. In the interest of obtaining more extensive data for analysis, the article on the nuclear program was selected for this study.

Figure 3.3. *Yahoo* Online Korean Dictionaries

야후! 사전

손님 [로그인] 야후! | 뉴스 | 게임 | 메일

전문자료검색 · 지식검색 · 교육 | 사전 홈

YAHOO! N2O YAHOO! 엔투오

야후! 사전

영한/한영 사전

검색

[영한/한영 사전으로 가기](#)

영영 사전

검색

[영영 사전으로 가기](#)

백과 사전

검색

☒ 전체 ☐ 제목 ☐ 이미지

[백과사전으로 가기](#)

IT 용어 사전

검색

☒ 전체검색 ☐ 제목검색 ☐ 내용검색

[IT 용어 사전으로 가기](#)

국어 사전

검색

[국어사전으로 가기](#)

경제 용어 사전

검색

[경제 용어 사전으로 가기](#)

3.3.3. Procedures

The individual participants had two meetings with the researcher—a preliminary training session and a second meeting in which paraphrasing and stimulated recall took place. At the beginning of the first meeting, ten minutes were devoted to distributing and answering a questionnaire requesting demographic information that included the length of residency in English speaking countries, experience with using a corpus and other reference tools, and use of computers and the Internet (see Appendix F).

At the preliminary meeting, the participants learned the basic text-searching functions of *MonoConc Pro* version 2.2. Based on the experience of the pilot study, this lasted approximately 30 minutes. This included loading a corpus, displaying frequency of occurrences, searching for single and multiple words, excluding words from consideration using wild cards, and displaying contexts. The subjects were asked to do practice activities using *MonoConc Pro* version 2.2 on a PC (personal computer). The activities included finding the most frequently occurring word in a corpus, identifying prepositions used with a certain verb, and deciding between meanings of synonyms such as “convince” and “persuade.” They were also asked to find various collocates of these words and their frequencies.

Next, they used the concordancing program for their own writing purposes for seven to ten days. This period of time was given to allow for participants to become accustomed to using the program. The extent to which each subject used the concordancing during this period varied. They were provided with a CD containing *MonoConc Pro* version 2.2 and a corpus consisting of *The New York Times* published from 2000 to 2002.

Fifteen minutes before the second meeting began, preparation for the computer screen recordings, loading the corpus to the concordancing program, opening the online dictionaries, and setting up the word processor margins, font and font size were completed. Also, the concordancing program and the computer screen recording were tested for their functionality. Five minutes after the task and before the recall session was spent setting up the playback and a second recording for the following recall session. Using a second computer, an audio recording of the recall session was also set up during this time because of the low quality of the audio recorder included in the screen recording program.

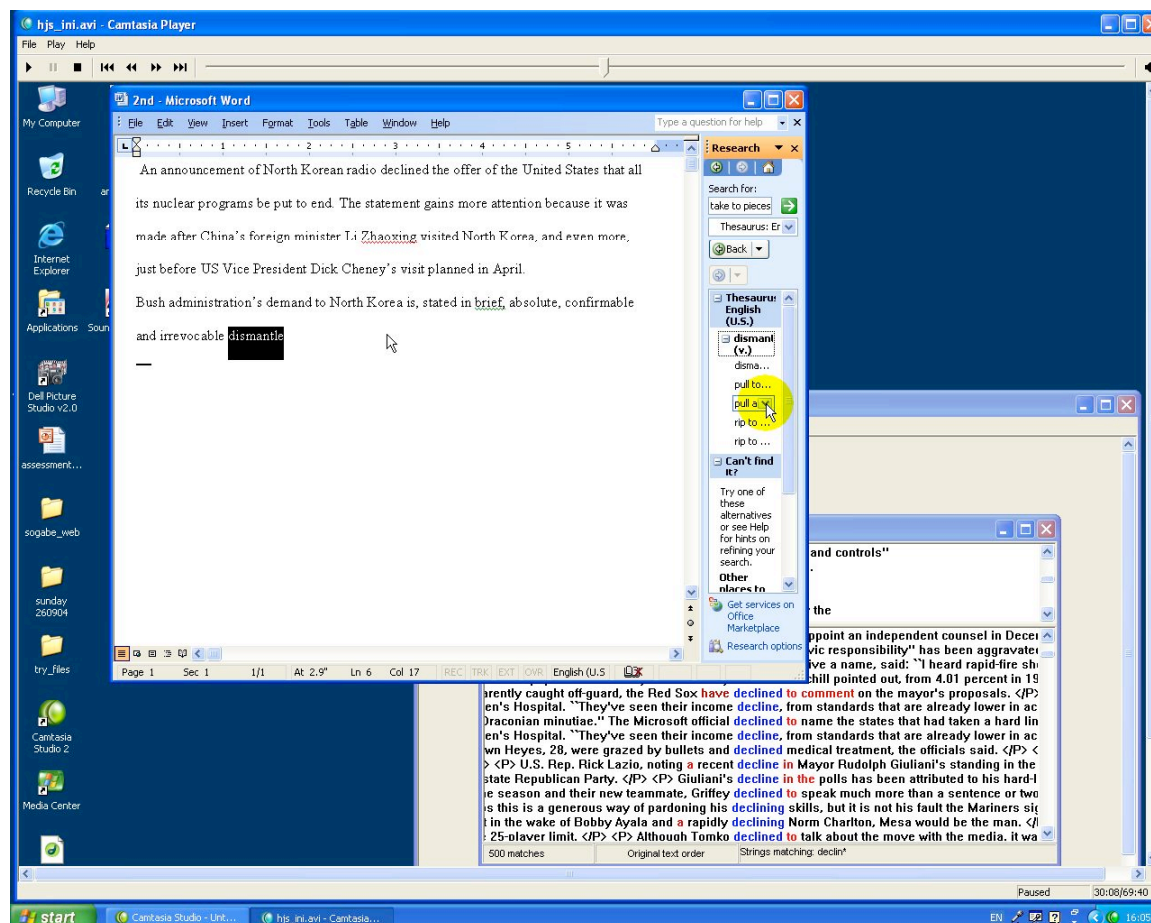
At the second meeting, the participants were given written plans for the paraphrasing session (see Appendix C) and an instruction sheet on how to paraphrase

(see Appendix D). They had five minutes to read these and ask questions and were then given the newspaper article. A part of the corpus consisting of *The New York Times* published from January 2000 to April 2000 had already been loaded on *MonoConc* before the session began. Subjects were not given a time limit and were told that the length of the paraphrasing should be half a page double-spaced, in 12-point Times New Roman font with 1" top and bottom margins and 1.25" left and right margins (approximately 150 words). The word processor was set to these defaults. After this explanation, the participants began to paraphrase.

While the participants were paraphrasing, all of the interactions on the screen, including *MonoConc Pro* version 2.2, Korean *Yahoo* Online dictionaries, and writing were recorded. The screen recordings were performed with *Camtasia Studio*, which records all interactions on a computer screen to a digital file (see Figure 3.4). This software did not interfere with the users' tasks during paraphrasing. The participants were notified at the beginning of the second meeting that this software was going to be used for the stimulated recall session that would be conducted immediately following the paraphrasing tasks.

Before the stimulated recall session, participants received an explanation of its purpose and procedure (see Appendix E) and had five minutes to read it and ask questions. Next, the participants met with the researcher individually and viewed their recorded screen interactions. When the researcher saw an interaction with the dictionary or concordancing program, he paused the video and asked the subjects what they had been thinking. The subjects were told that they could also pause the video at anytime and tell the researcher what they had been thinking. All of the conversations were in Korean and recorded digitally on the computer along with a copy of the original screen capture of the paraphrasing activity. Relevant portions of the conversation, translated into English, can be found in the results section.

Figure 3.4. Screen recording for stimulated recall



In the pilot study, the subjects spent one hour for paraphrasing, while the recall session took an average of 48 minutes. These recall sessions included researcher-participant interaction and replayed support cues. In order to ensure smooth scheduling, 60 to 90 minutes for the task was estimated. Since there was no time limit for paraphrasing, the initial data collection time varied. A five-minute break following the task was given to avoid fatiguing the subjects. The stimulated recall session lasted approximately one hour. This allowed enough time for interactions between the researcher and each subject including replay of the support cues.

Table 3.4. Time frame of the meetings

Meeting	Purpose	Time spent (minutes)
First Meeting	Demographic Data	10
	<i>MonoConc</i> Tutorial	30
	Total (Meeting 1): 40	
	7-day interval between the two meetings	
Second Meeting	Preparation	15
	Information Sheet for paraphrasing	5
	Paraphrasing	60 – 90
	Break & Preparation	5
	Information Sheet for SR	5
	Stimulated Recall	60
	Survey	15
	Total (Meeting 2): 165 – 195	
Total (Meeting 1+ Meeting 2): 205 – 235		

After finishing the stimulated recall session, 10 additional minutes were taken for the subjects to complete a survey for the purpose of understanding their attitudes toward using a corpus as a reference tool for writing and subjects' personal use of the corpus for the past week. (See Table 3.4 for the time frame of the meetings.) The questions focused on the subjects' opinions on the usefulness of the corpus, ease of operating the concordancing program, and effects on the subjects' writing (see Appendix G).

3.4. Analysis

3.4.1. Interrater reliability

In order to ensure the objectivity of the raters, an interrater reliability test was carried out. This test ensured that what the researcher saw in the data could also be seen by the other rater, and that the researcher's ratings on each episode are consistent. There

were two raters: the researcher and another person who was not a participant in the recall session. This third-party rater (TPR) was trained in the categories for rating. The TPR was given information about the procedure of the stimulated recall session and events that were being recalled. For purpose of training, the TPR was given a sample set of data from the pilot study with a coding sheet. The researcher then explained the meaning of each code and events to which each code applied. The TPR coded the sample data and, when the coding was done, the codes were compared to that of the researcher. When there was any difference in coding, the researcher asked the reason for the difference in coding and compared the reasons why they should be coded differently. The TPR was given another example set of data to practice coding. The coding was compared again with that of the researcher. The TPR was also provided with explanations of the coding scheme if there were any differences.

After the training phase, the TPR was given the entire set of data that was the recorded screen interactions during the stimulated recall session with audio utterances and length of time for each episode and event. In order for the TPR to code conveniently, an HTML page with utterances was divided into episodes (see next section) and QuickTime movies for individual episodes were embedded in the page. The TPR then watched the movies and read the utterances on the page for coding. A link to the reference for coding was present on the page. This page provided radio buttons for a code that the TPR needed to choose from for multiple codes and text areas for additional information. When the coding was done, the TPR saved the HTML page and sent it to the researcher via formmail at the end of each episode. The researcher also used this HTML page for coding and compared the coding. There were differences between the coding done by the researcher and the TPR. The researcher and TPR then had a meeting in order to look closely at the differences. During the meeting, they explained the reasons for their coding to each other and tried to decide on the code that is most appropriate.

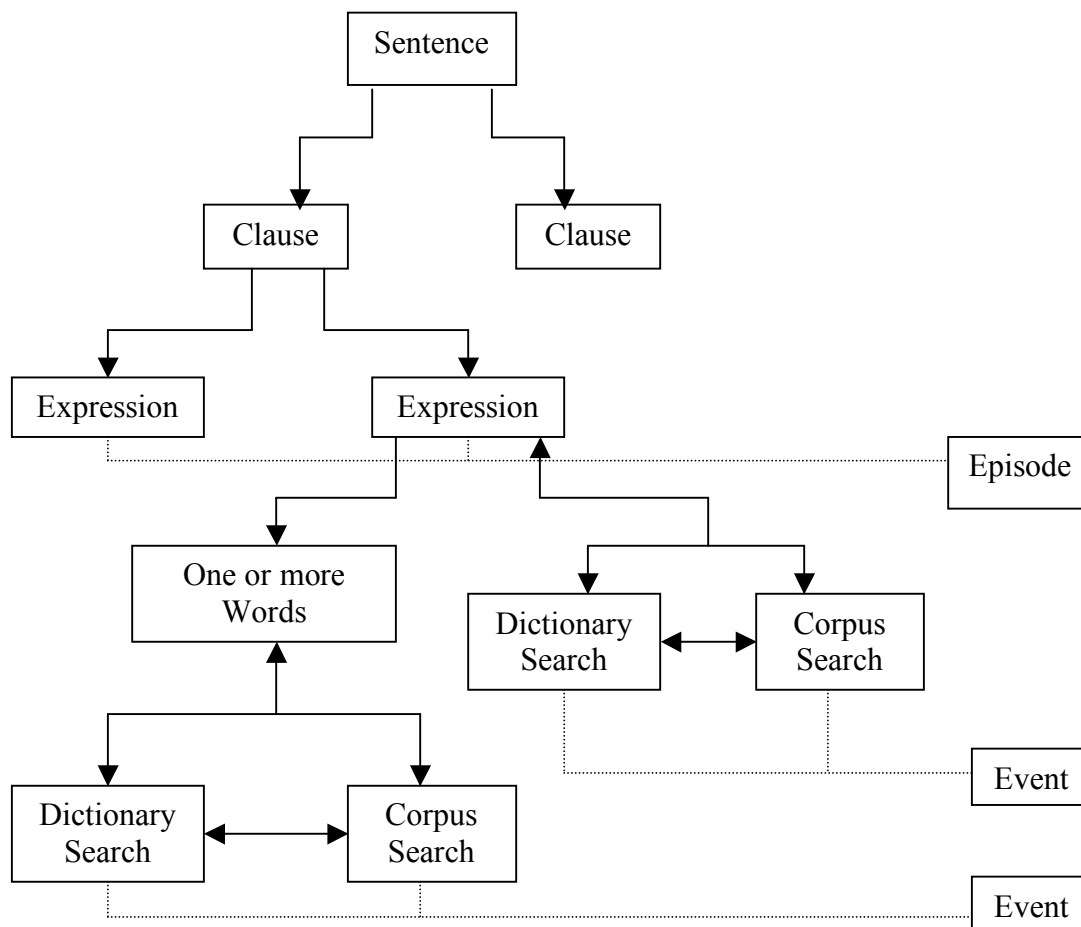
3.4.2. Units of analysis

The unit of analysis for this study consists of a single written “expression” from the subject’s paraphrase. An expression in this case may be one or more words (see Figure 3.5) that form a meaningful unit. For example, when the subjects search in dictionaries or a corpus, their search can consist of a single word, multiple words, or a clause. One expression can involve a single search or a sequence of various reference tools.

The reason the expression can be one word or multiple words is that in some cases, a search word in Korean could be a phrase or even a clause in English or vice versa. One expression represents one episode. An episode can then be a set of processes of constructing one or more words. This set can include one or more searches in reference tools. Each search is counted as one event. An event is a single use of any reference tool.

It was clear that the subject was working on one expression (one unit) although he/she might be working with different references and search words. So long as the subject continued to work on one expression, all the interactions involved in the expression were considered within the same analytical unit or expression. For paraphrasing, for example, if the subjects desired to use a different word from the original text and did not know the English equivalent, they could use a Korean dictionary and get an English translation. With an English translation they referred back to a Korean dictionary to verify the meaning or used *MonoConc* in order to search for examples or collocations. In this case, the sequence of searches was included in the same unit.

Figure 3.5. Units of analysis



3.4.3. Coding

A coding scheme was developed based on the research questions. The coding scheme addressed the tool being used, reason for using each tool, information obtained, strategy used, ways the corpus complements dictionary use, and effects on writing. In order to fine-tune the coding, a top-down coding scheme was developed on the basis of the research questions. The initial coding scheme was somewhat similar to that of the pilot study (see Appendix I). This coding consisted of three parts: initial search actions,

search processes, and output. The initial search actions included use of reference tools, choice of search words (or phrases) including parts of speech, and reasons for the search. The coding for search processes included information the subjects looked for, information they found, the decisions on whether or not to use the information they found, the reasons for the decisions, and the strategy they implemented during the search. The output part included the final expressions that the subjects wrote in their writings through the initial search actions and search processes. After this initial coding process, the final version of the coding scheme was then determined.

The analysis of data obtained from the screen recording and stimulated recall was quantitative and qualitative. Data that were categorized and quantified were then tested for significance and illustrated through charts and other media as needed. For example, reasons for consulting a corpus, strategies used, types of information searched for, including parts of speech and ways of complementing dictionaries, categorized and counted as part of a quantitative analysis. Additionally, qualitative analysis was carried out on data that showed patterns of using reference tools and other strategies, detailed reasons for successful (or unsuccessful) searches, cases of consulting a corpus in addition to dictionary searches, cases of consulting a corpus because of the lack of information in dictionaries, and the effects of using a corpus on writing or self correction. Lastly, the data on subjects' attitudes towards using a corpus gathered from the surveys was described.

The following paragraphs show how research data were obtained and what the anticipated results were for each research question:

- (1) What are the reasons that ESL learners consult reference tools while writing in English?

Research data for this question were obtained during the stimulated recall session in response to interview questions such as "What was the problem you needed to solve when you decided to use one of the reference tools?" Anticipated

possible reasons included: uncertainty about existence of an expression; need to study or verify the meaning of an expression in context; need to know appropriate usage of the expression in context.

- (2) What kind of information do learners look for in a corpus and dictionaries while writing in English?

Research data for this question were obtained from the screen recording and during the stimulated recall session. The screen recording showed, in part, what the subjects were reading in dictionaries and a corpus. The observable interactions included words the subjects searched for, clickable words or phrases they used in dictionaries and a thesaurus. The screen recording also showed what words the subjects looked for and what kind of information they looked at (e.g., collocation frequency, more context). There were times when it was not clear what the subjects were reading or looking for. In this case, the subjects were asked a question such as “What were you reading in the dictionary when you clicked on this word?” Anticipated possible kinds of information included: search words or phrases, parts of speech, definitions, synonyms, words or phrases in context.

- (3) What strategies do learners use when they consult a corpus?

Research data for this question were obtained from the screen recording and during the stimulated recall session. The screen recording showed, in part, what strategies the subjects choose to use for better results. Strategic behaviors that could not be observed on the screen recording were elicited during the stimulated recall. Research data were obtained in response to interview questions such as “What were you thinking when you changed the search word?” The anticipated possible strategies included: using a combination of reference tools; searching for words coming before and after; rearranging the search words; using less specific search words; using more specific search words; using fewer search

words; paying attention to the context before and after; checking spelling; looking for context-specific words; paying attention to the part of speech of the search word; seeing collocation frequency; using wildcards for searches (e.g., * and @); finding out which meaning is used for multiple-meaning words; paying attention to the parts of speech of collocates; paying attention to the sentence structures in the results; remembering results from previous searches; changing the part of speech to accommodate the example; looking for examples in specific context; paying attention to the number of results; combining dictionary search; looking for modifiers for noun search; finding example in a similar context; canceling search when getting enough results to save time; using advanced collocation frequency options.

(4) How does a corpus complement the use of dictionaries?

Research data for this question were obtained during the stimulated recall session in response to interview questions such as "Why did you use the same search word in *MonoConc* after using the English-English dictionary?"

Anticipated possible ways of complementing dictionary use included: needing to know more collocations other than those shown in the dictionary; needing to see the use of the search word in a specific context that is not found in dictionaries; needing to get examples of using a phrase in a sentence quickly; needing to see more examples in order to understand the meaning of a word and grammatical structure in context.

(5) What are the effects of using a corpus on the accuracy of learners' writing?

Research data for this question were obtained from two third-party English-native raters' judgment on the subjects' writings. The judgment was made whether an expression was native-like or not. The two raters were American English native speakers who were not involved in any procedures of this research. These raters graded the writings separately and reconciled any

differences in their judgments if needed. Based on these grades, the results showed the ratio of successful and unsuccessful episodes, sequence of consultation, results found in a corpus, categories of errors, information (or sequence) that causes nonnative- or native-like expressions. These results then lead to conclusions for training to solve problems caused by corpus use.

(6) What are learners' attitudes toward using a corpus as a reference tool?

Research data for this question were obtained from the survey. The survey contained questions that focus on effectiveness, helpfulness, and ease of using *MonoConc* for the subjects' writing.

CHAPTER IV

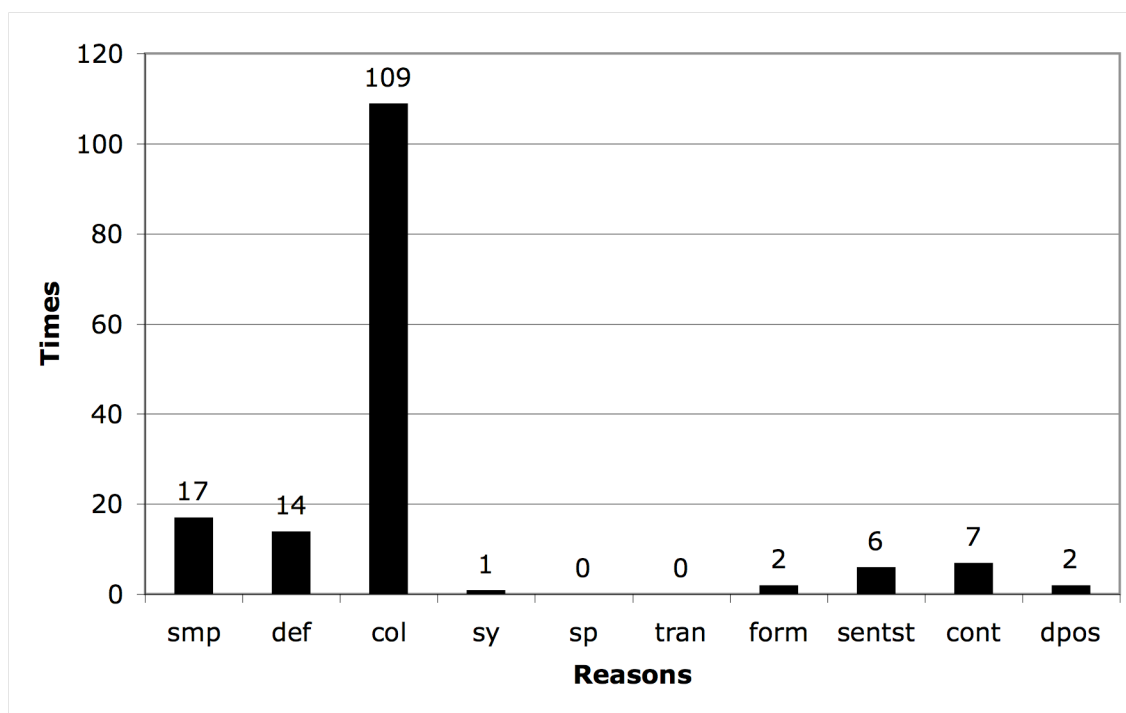
RESULTS

This chapter presents the results of this study in detail according to the six research questions previously described. The questions are discussed first, followed by tables and examples from the stimulated recall sessions. The tables include line numbers that indicate speech turns of the researcher and participants. The tables also include observations of screen-captured video and actions taken by the subjects during stimulated recall.

4.1. What are the reasons why ESL learners consult a concordancing program while writing in English?

The subjects consulted the concordancing program in order to obtain several kinds of information, including collocations, definitions of search words, clarification of their writing, context, and parts of speech. The reason for using dictionaries was to obtain definitions and synonyms of search words. Dictionaries are sometimes also used to look for sample sentences. When they were not able to find the sample sentence they were looking for, subjects would then use *MonoConc* for more sample sentences. Occasionally subjects also used dictionaries to find the part of speech of a word. Thesauri were solely used to look for synonyms. In the remainder of the chapter, samples from each category of the data will be discussed.

The ten subjects had a choice of five tools during this study. The first tool discussed here is *MonoConc*. The most common reason for consulting *MonoConc* was collocation (109 times, 69.0%, see Figure 4.1).

Figure 4.1. Reasons for using *MonoConc*

smp: sample, def: definition, col: collocation, sy: synonym, sp: spelling, tran: translation, form: word form, sentst: sentence structure, cont: context, dpos: defining part of speech

In Table 4.1, the participant typed ‘accus*’ in *MonoConc*. She wanted to see what came after ‘accuse’ in object position. She needed to know how to use a preposition and object with this word and whether or not a preposition was actually used with ‘accuse.’ The subject then saw an example with an object following ‘accuse,’ and wanted to search further about the possibility of a preposition following the word (see line 2). The results showed that this participant had been looking for the difference in meaning between ‘accuse + of’ and ‘accuse + object.’

Table 4.1. Subject JWK using *MonoConc* with search term "accus"

No.	Observation	Conversant	Verbatim
1	VO Typing 'accus*' in M	Researcher	Why did you look for 'accus*' in M?
2		JWK	I was looking for what came after it as an object of 'accuse.' I wanted to know how to use a preposition and object or if an object came without a preposition
3	AC Pointing at the article	JWK	Here, an object came after 'accuse,' so I was looking for this kind of example and whether or not a preposition was used after it.
4		Researcher	What were you looking at?
5	AC Pointing at the results		I was looking for the difference in meaning between 'accuse+of' and 'accuse+object.'
Original sentence in the article (OS): It used typically unrestrained language in accusing the United States of secretly planning a war.			
Resulting sentence in paraphrase (RS): It utilized generally unlimited language in charging the US of secretly planning a war.			

VO: Behaviors that appeared in the screen recording

AC: Behaviors of subjects while watching the screen recording during the stimulated recall sessions

M: *MonoConc*, EE: English-English dictionary, EK: English-Korean dictionary, KE: Korean-English dictionary, T: thesaurus, W: word processing program

A very distant second was samples with only 17 searches (10.8%), followed by searches for definitions (14 times, 8.9%). One of the things that *MonoConc* provides is a variety of sample sentences. *MonoConc* provides many more samples than dictionaries or thesauri. When subjects look for samples, they usually know the meaning of the search words or search phrase. They want to see if such a word or expression exists in a real life situation, so that they are not creating a nongrammatical statement. Subject HJS took advantage of this resource (see table 4.2). Dictionaries provide examples of a single word or phrase. This is not always convenient, because it requires more search time. *MonoConc* enables users to target a desired word or phrase quickly.

Additionally, it is much harder to find examples of phrases in a dictionary, in contrast with the convenience of typing in a phrase in *MonoConc* and finding related search results.

Table 4.2. Subject HJS using *MonoConc* with search term "destruction @ country"

No.	Observation	Conversant	Verbatim
1	VO Typing 'destruction @ country' in M	HJS	I was wondering if I could use it like this. I made it up, so I was checking to see if it was used in real life.
2	VO Typing 'the destruction of the country' in W		
3	VO M showing one example (destruction of his own country) at the article	Researcher	There weren't many examples. What did you decide?
4		HJS	It wasn't used very frequently, but I thought it was possible to use it to convey the meaning.
5	VO Typing 'destruction of the country' in W		
OS: "Irreversible nuclear dismantling" is nothing other than a noose to stifle us after eradicating our peaceful nuclear-energy industry," it said.			
RS: North Korea expressed its aversion to this demand, saying that US aims to spy on its military facilities, agitate a war against North Korea and furthermore, put them trapped in the destruction of the country.			

In Table 4.2, the subject began with the phrase 'destruction @ country.' She wanted to see if this phrase existed in real life so that she could use it in her writing. Although she found only one example, it was enough to convince her that the phrase she chose conveyed the correct meaning (see line 4).

Thus, students found the definitions by understanding the context of examples provided by *MonoConc*. Relationships between the search words and surrounding words were reviewed. This task required the users to understand most of the words used in the samples. Because of this, they needed to have an adequate level of English proficiency.

Table 4.3. Subject JWK using *MonoConc* with search term "restrain"

No.	Observation	Conversant	Verbatim
1	VO Typing 'restrain' in M	JWK	I wanted to find out the meaning through examples. I knew the meaning and wanted to see how it was used. What I thought was right.
2		Researcher	What were you reading?
3		JWK	I was reading some sentences to see how to use it. I also looked for the meaning of 'restrain.'
4		Researcher	What did you learn?
5			I found out the meaning of 'restrain' and changed it to a different word.
OS: It used typically unrestrained language in accusing the United States of secretly planning a war.			
RS: It utilized generally unlimited language in charging the US of secretly planning a war.			

In Table 4.3, this subject was attempting to replace 'restrain.' Before she replaced 'restrain,' she needed to know its meaning. She read some sample sentences to see how her search word was used. She discovered the meaning of 'restrain' and changed it to a different word.

Subjects also used *MonoConc* for the purposes of clarification. Clarification is defined as a user conducting a search to see whether their writing was correct or not. In Table 4.4, the participant typed 'planned visit' in *MonoConc*. He noticed that in the article, the phrase used was 'that is planned for April.' He wanted to shorten it to 'planned visit,' but wanted to see some examples first. In *MonoConc*, he saw 'visit to + country by + person,' and then decided to change the structure of his sentence.

When the reason for using *MonoConc* is to look for context, the surrounding words are what the users usually look for. Subjects looked for the situation the particular surrounding words were used in and whether they applied to the current situation.

Table 4.4. Subject MKS using *MonoConc* with search term "planned visit"

No.	Observation	Conversant	Verbatim
1	VO Typing 'planned visit' in M		
2	AC Reading the article	MKS	Here, it says, 'that is planned for April.' I was going to shorten it and write 'planned visit.' I think it was helpful that I saw the examples. I was going to write 'planned visit of Dick Cheney to North Korea.' But I saw in M 'visit to + country by + person.' So, I changed the structure of the sentence. I came up with 'planned visit,' so I wrote it first before I looked for it in M. But I thought it sounded awkward, so I wanted to be sure. I then saw the expression in M and decided to use it in W.
OS: The statement carried by Radio Pyongyang and monitored by news agencies in South Korea came just after a visit to North Korea by China's foreign minister, Li Zhaoxing, and shortly before a visit to the region by Vice President Dick Cheney that is planned for April.			
RS: The statement through Radio Pyongyang which was monitored by South Korean news agencies appeared just after China foreign minister, Li Zhaoxing's visit to North Korea and shortly before a planned visit to the region by the US Vice-President Dick Cheney in coming April.			

Table 4.5. Subject HJS using *MonoConc* with search term "decline"

No.	Observation	Conversant	Verbatim
1	VO Copying 'decline' and pasting 'decline' into M	HJS	I was wondering if I could use 'decline.'
2		Researcher	Why did you look for 'decline'?
3			To see if I could use it in this context, and to see the part of speech.
4		Researcher	What kind of context were you looking for?
5		HJS	For example, I wanted to see whether it was used as a noun or verb. And I wanted to see if 'decline' could be used in a political context
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: An announcement of North Korean radio declined the offer of the United States that North Korea should stop all its nuclear programs.			

In Table 4.5, the participant wanted to use ‘decline’ in a political context. He found some examples in *MonoConc* and was checking to see if they were used in the same context.

MonoConc was not the only tool the subjects used during this study. The thesaurus yielded 45 searches for synonym or word change purposes.

Table 4.6. Subject MAP using thesaurus with search term "rejection"

No.	Observation	Conversant	Verbatim
1	VO Typing 'rejection' in T	Researcher	Why did you look for 'rejection' in T?
2		MAP	I wanted to change 'reject' to a different word.
3	AC Pointing at results in T	Researcher	What were you looking at?
4		MAP	I thought about 'denial' (in the results), but it didn't seem right. So, I decided to use 'rejection.'
5		Researcher	Verb or noun?
6		MAP	Noun.
7	VO Typing 'official rejection on' in W		
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: Radio Pyongyang announced the rejection right after Li Zhaoxing, China's foreign minister, visited to North Korea.			

Table 4.6 shows that the participant used the thesaurus to search for ‘rejection.’ She wanted to use a different word instead of ‘reject.’ The results showed ‘denial,’ but this did not seem like the correct choice to the participant. In the end, she decided to use ‘rejection,’ but did not explicitly say why.

The English-Korean dictionary was used 26 times for definitions and 12 times for synonyms or word changes, with a lower number of other uses. This corroborates Laufer

(1993), who found that definitions plus examples gave more help than using either of these resources on their own.

Table 4.7. Subject KWA using English-Korean dictionary with search term "inconclusive"

No.	Observation	Conversant	Verbatim
1	VO Typing 'inconclusive' in EK	Researcher	You're looking for 'inconclusive' in EK.
2	AC Reading the article	KWA	I wasn't sure about the meaning of 'inconclusive.' I looked for it in a dictionary to get the definition.
3		Researcher	Did you get the correct definition?
4		KWA	Yes.
OS: But in the latest statement, it appears to be setting the stage for another inconclusive effort.			

The search displayed in Table 4.7 shows that the participant used the English-Korean dictionary. The participant conducting this search was looking for 'inconclusive.' He was not sure about the meaning, so he used the dictionary to get a definition and thought that this was the correct definition. Similarly, Nist and Olejnik (1995) found that the quality of a dictionary definition was the determining factor in the quality of learning and believe that the dictionary can be a substantial contributor to the vocabulary learning process.

The majority of the searches in the Korean-English dictionary (23 times) were for translations, followed by 12 searches for synonyms. In Table 4.8, the subject was trying to replace 'talk' and search for the English equivalent of the Korean word, 'dedam' (discussion) in the Korean-English dictionary. She found 'conversation' in the dictionary and used it in her writing.

Table 4.8. Subject KYL using Korean-English dictionary with search term "dedam"

No.	Observation	Conversant	Verbatim
1	VO Typing 'dedam' (discussion) in KE	Researcher	You were typing 'dedam' in KE.
2		KYL	I was going to change 'talk,' so I was looking for other words to replace 'talk.' 'Talk' seems to be similar to 'dedam.' I didn't find anything I liked. I thought about 'discussion,' but it's not what I wanted.
3	AC Pointing at 'conversation' in KE	KYL	I thought I saw 'conversation.' Here, I changed it to 'conversation.'
4	VO Typing 'conversation' in W		
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: North Korean radio announced on Saturday that they rejected the formula the United State suggested as its pragmatic position in conversation which is aimed to end North Korea's nuclear programs.			

Table 4.9. Subject YJK using English-English dictionary with search term "noose" and "accord"

No.	Observation	Conversant	Verbatim
1	VO Typing 'noose' in EE	Researcher	What were you thinking when you searched for 'noose' in EE?
2		YJK	I didn't know the meaning.
3		Researcher	What were you looking at?
4		YJK	I wanted to know what the meaning was in the sentence. I had to understand it in order to paraphrase. Whenever I used the <i>Yahoo</i> dictionary, I was looking for meanings.
5	VO Typing 'accord' in EE	Researcher	You were also looking for 'accord' in EE?
6		YJK	I didn't know the meaning.
OS: "Irreversible nuclear dismantling" is nothing other than a noose to stifle us after eradicating our peaceful nuclear-energy industry," it said.			
RS: North Korea needs to keep part of nuclear program for civilian uses and America's obstinate pressure on the total eradication of nuclear program is to take advantage on the war to destroy North Korea.			

The tool used least was the English-English dictionary. This tool was used mostly for obtaining definitions and looking for synonyms or word changes. Because this type of dictionary provides both definitions and synonyms, it was used exclusively for those two reasons, while the English-Korean dictionary was used for a variety of reasons including definitions, synonyms, information about parts of speech, and examples.

The subject in Table 4.9 was searching for ‘noose’ in the English-English dictionary because she did not know the meaning. She needed to know the meaning in order to paraphrase. She also looked for ‘accord’ in the English-English dictionary because she did not know the meaning of it either.

Table 4.10. Subject KYL using English-English dictionary with search term "bottom-line"

No.	Observation	Conversant	Verbatim
1	VO Typing 'bottom-line' in EE	Researcher	There is a noun and adjective, and you chose an adjective.
2		KYL	Here, it is used as an adjective, so I was looking for an adjective. I saw 'pragmatic,' but it was different from what I thought. I thought it was awkward.
3	VO Typing 'pragmatic position' in W	Researcher	So, what did you do?
4		KYL	I used 'pragmatic' instead of 'bottom line.'
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: North Korean radio announced on Saturday that they rejected the formula the United State suggested as its pragmatic position in conversation which is aimed to end North Korea's nuclear programs.			

Table 4.10 shows that ‘bottom line’ was the search phrase used in the English-English dictionary. She saw that one word was an adjective and another word was a noun. She was looking for an adjective and saw the word ‘pragmatic.’ This was enough

information for her to change her mind and use ‘pragmatic’ instead of ‘bottom line’ (see line 2). For this participant, comparing words based on the part of speech helped her reach a decision.

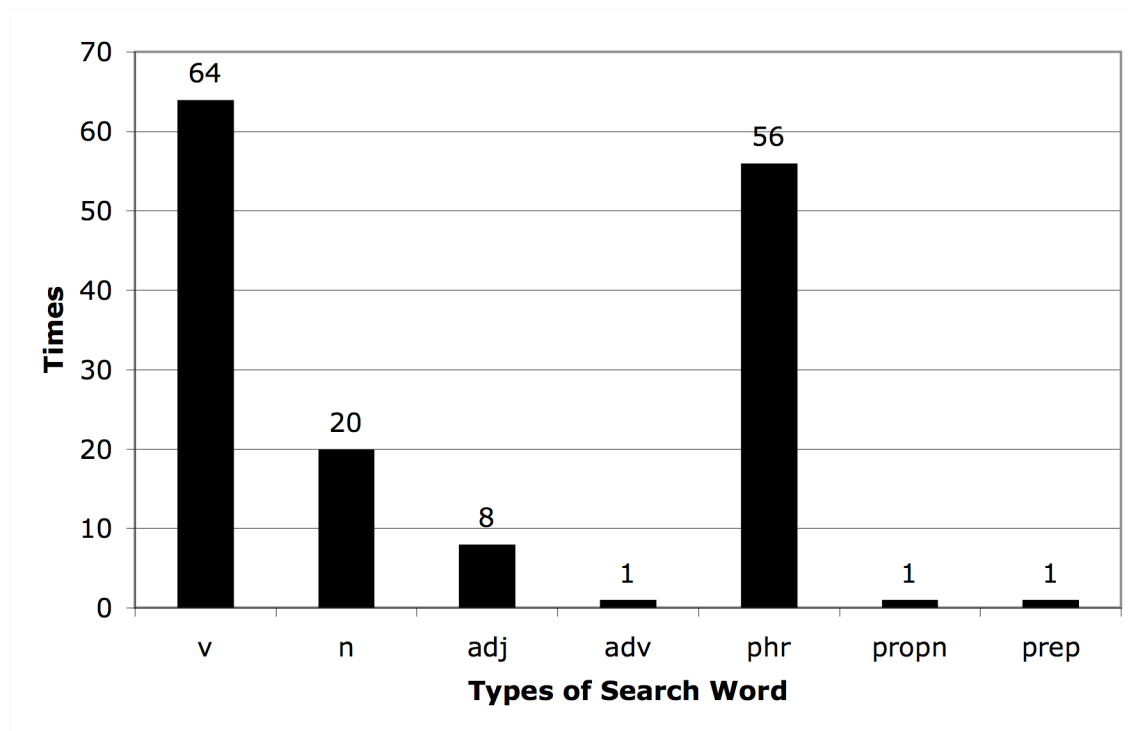
As demonstrated in section one, the majority of the searches are performed to look for collocation, followed by sample sentences and definitions. Most of the collocations participants looked for were words or phrases following a verb. By looking at the context of sample sentences in *MonoConc*, they came up with the meaning. Sample sentences also served as resources that clarified whether the words and phrases they were writing were correct or not. When subjects used *MonoConc* and the dictionary in order to define the parts of speech, they looked for surrounding words and tried to define the parts of speech accordingly. When the dictionary was used, as expected, it was mostly for definitions. When subjects wanted to look for synonyms of words, they used both dictionaries and thesauri.

4.2. What kind of information do learners look for in a corpus and dictionaries while writing in English?

Research data for this question were obtained from the screen recording and during the stimulated recall session. The screen recording showed, in part, what the subjects were reading in dictionaries and a corpus. The researcher was able to observe words the subjects searched for. The screen recording also showed the selections the subjects made when looking for definitions and examples in the dictionaries and thesaurus. In the screen recording the researcher was also able to see what words the subjects looked for and what kind of information they were looking at (e.g., collocation frequency and more context). Based on experience from the pilot study, there were times when it was not clear what the subjects were reading or looking for. When this occurred, the subjects would be asked a question such as “What were you reading in the dictionary when you clicked on this word?” Anticipated possible kinds of information include:

search words or phrases, parts of speech, definitions, synonyms, words or phrases in context.

Figure 4.2. Types of search word

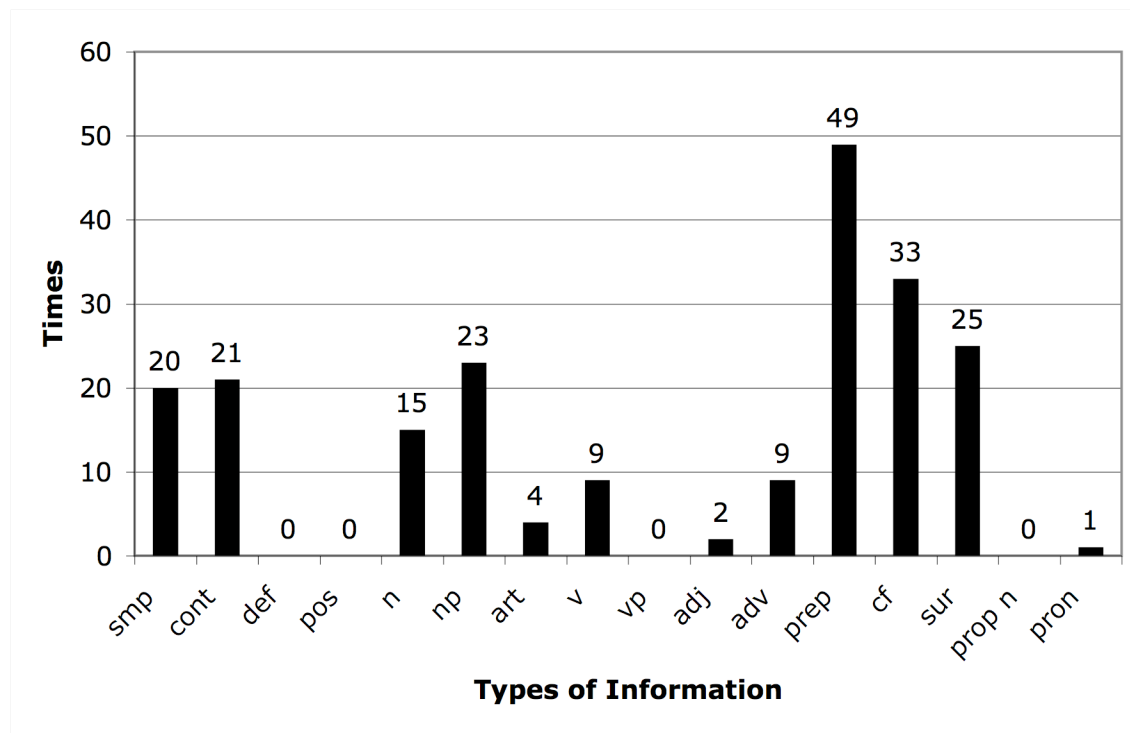


n: noun, v: verb, adj: adjective, adv: adverb, phr: phrase, propn: proper noun, prep: preposition

The participants had five sources for entering in and finding information about search words. The most commonly used tool was *MonoConc*. One hundred and fifty-two searches (47.4%) were made with *MonoConc*. The second most used was the English-Korean dictionary (62 times, 19.3%). The least used was the English-English dictionary (15 times, 4.7%). The most common part of speech or “function” of the search words in *MonoConc* was the verb (64 times, 42.4%, see Figure 4.2). Phrases were used as search words 56 times (37.1%), and the least used were search words that were proper

nouns and prepositions (one search each). In the following section, examples of each tool will be shown, along with detailed interactions between the subjects and *MonoConc*.

Figure 4.3. Information looked for in *MonoConc*



smp: sample, def: definition, cont: context, pos: part of speech, n: noun, v: verb, adj: adjective, adv: adverb, phr: phrase, prop n: proper noun, pron: pronoun, prep: preposition

The information most frequently searched for in *MonoConc* related to prepositions (49 searches, 23.2%, see Figure 4.3), with collocation frequency being the purpose of 33 searches (15.6%). The first example shows a subject searching for a preposition that followed a verb (see Table 4.11). Her first choice was to enter ‘skeptical about’ in *MonoConc* to make sure it was correct. While she was not sure if ‘skeptical about’ was correct, she also entered ‘skeptical on’ in *MonoConc*. When she did not see any results, she decided that ‘skeptical about’ was the best answer (see line 6).

Table 4.11. Subject JYP using *MonoConc* with search term "skeptical about"

No.	Observation	Conversant	Verbatim
1	VO Typing 'skeptical about' in M	Researcher	Why did you look for 'skeptical about' in M?
2		JYP	I wrote 'skeptical about,' and was making sure if it's correct. I was also looking to see if 'skeptical on' showed up in M. I thought I was looking for 'on' also.
3		Researcher	What were you looking at?
4		JYP	I wanted to see how often it's used in sentences and if the context was similar to that of the sentence I was writing.
5	VO Typing 'skeptical on' in M	Researcher	This time, you were looking for 'skeptical on' in M.
6	VO M is searching for examples. VO Canceling the search.	JYP	If it didn't show any results within this much time, there wouldn't be any results. So, I used 'about.' I looked for both in M and found more examples with 'about.'
OS: . . . ending North Korea's nuclear programs, raising doubts about whether the fitful negotiations are making even limited progress.			
RS: Because of this announcement, people are skeptical about the effectiveness of this negotiation.			

The participant in Table 4.12 was unsure if 'agree with' or 'agree to' was correct. She wanted to use the phrase 'agreed to the American terms' (see line 4). In one example, she saw that a noun came after 'to.' In other examples, she saw 'to' used as an infinitive (see line 6). After this, she returned to *MonoConc* for more examples. She found that many examples included 'to + noun,' but most examples were composed of 'to + verb.' She saw that some even used the formula of 'agreed with + noun.' She decided to change her answer to 'agreed to + verb' after seeing the examples because it was used frequently.

Table 4.12. Subject JDO using *MonoConc* with search term "agreed"

No.	Observation	Conversant	Verbatim
1	VO Typing 'agreed' in M	Researcher	Agreed?
2		JDO	I wasn't sure whether it's 'agree with' or 'agree to.' I wanted to write 'agreed' instead of 'accepted.'
3	VO Reading a sentence in M	Researcher	What were you thinking when you read 'agree to'?
4		JDO	It says, 'accept American terms.' I wanted to use 'agreed to the American terms.' I paraphrased 'common negotiation goal' to 'American terms.' So, I could write 'agreed to a plan,' 'agreed to an American plan.'
5		Researcher	Other examples also contain 'to.' Why did you choose this example?
6		JDO	In this example, a noun comes after 'to.' Others show 'to + verb.' 'To' is used as an infinitive. The 'to' in this example is a preposition, and other 'to's' in other examples are infinitives. I wanted to use 'to + noun.'
7	VO Typing 'agreed to the American terms' in W. VO Reading more examples in M	Researcher	Why did you go back to M?
8		JDO	Because there were many examples with 'to + noun.' Most of them have 'to + verb,' and some with 'agreed with + noun.' I thought I might be able to use 'agreed with,' but 'agree with' is followed by a person. In my case, you and I agree on American terms.
9		Researcher	After you used M, you changed it to 'agreed to + verb.'
10		JDO	Yes, I changed it to a verb.
11		Researcher	Why did you change it to a verb after looking for examples with 'agree to + noun'?
12		JDO	I wanted to use something that was used frequently. In M, most of them are used with 'to + verb' form. For example, it's correct to write 'agreed on the American terms' or 'agreed to accept the terms.' But I noticed that there were more examples with 'agree to + verb' than those with 'agreed to something.' So, I used more frequent forms.

Table 4.12. Continued

OS: Bush administration officials have said that North Korean negotiators discussed accepting the American terms as a common negotiating goal in the last round of talks, held in Beijing in February.

RS: According to Bush Administration officials, North Korea has agreed to accept the American terms as a common negotiating goal in the last round of talked, held in Beijing in February.

In the thesaurus, the most common information searched for were verbs (19 times). In the episode shown in Table 4.13, the participant typed ‘dismantle’ in the thesaurus. She was thinking that she could possibly use ‘deconstruction.’ She saw that ‘demolish’ and ‘destroy’ mean to ‘totally destroy’ in Korean. She thought that ‘dismantle’ was the most appropriate choice, because all of the words mean to destroy.

Table 4.13. Subject MKS using thesaurus with search term "dismantle"

No.	Observation	Conversant	Verbatim
1	VO Typing 'dismantle' in T	MKS	'Dismantle' means 'dismantle' (in Korean). I was wondering if I could use 'deconstruction.' 'Demolish' and 'destroy' mean 'totally destroy' (in Korean). All the words mean to destroy, so I thought 'dismantle' was the most appropriate.
2		Researcher	What else did you see?
3		MKS	Nothing. I thought that 'dismantle' was appropriate for this context.
OS: The statement rejected the American demand for a "complete, verifiable, irreversible dismantling" of the country's nuclear programs.			
RS: North Korea in the statement adamantly turned down the US demand for a “entire, verifiable, and irrevocable dismantling” of the North Korea’s nuclear programs.			

When participants used the English-Korean dictionary, the most common reason was to find a definition (23 times, 40.4%). Table 4.14 shows the participant who was looking for ‘move’ in the English-Korean dictionary. She knew it was a noun in the

news article and was looking to see the meaning of the noun form. There are no results shown in the table, as the participant was only searching for some information.

Table 4.14. Subject MAP using English-Korean dictionary with search term "move"

No.	Observation	Conversant	Verbatim
1	VO Typing 'move' in EK	Researcher	You were looking for 'move' in EK.
2		MAP	I wasn't sure about the meaning because 'move' was used as a noun in the news article. I was looking to see the meaning of the noun form.
OS: The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. . . .			

Verbs were the most common information looked for in the Korean-English dictionary (15 times, 42.9%).

Table 4.15. Subject KWA using Korean-English with search term "palsang"

No.	Observation	Conversant	Verbatim
1	VO Typing 'palsang' (occur) in KE	Researcher	You're typing 'palsang' (occur) in KE.
2		KWA	North Korea is criticizing the US policy because the peninsula is in danger. This could cause a nuclear war. I was looking for something like this. So, I didn't use 'led.' I found 'generate' in KE, but it was a little awkward. So, I tried to use 'occur.'
3		Researcher	Occurrence?
4		KWA	Yes. I used 'occur.'
OS: The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime . . .			
RS: A North Korea's official said that the present policy of the U.S. just makes the Korean peninsula in danger, which can occur a nuclear war.			

The participant shown in Table 4.15 typed ‘palsang’ (occur) in the Korean-English dictionary. He was talking about North Korea criticizing U.S. policy. He did not use ‘led.’ He found ‘generate’ in the Korean-English dictionary, but thought it sounded awkward. He then decided to use ‘occur.’

The English-English dictionary was rarely used, but when it was, the most frequent reason was for a definition (seven times, 50.0%). In Table 4.16, the subject was looking for ‘explicitly’ in the English-English dictionary. She was looking for the definition and was successful in finding it. However, her resulting sentence was awkward, despite the information provided in *MonoConc*.

Table 4.16. Subject JWK using English-English dictionary with search term "explicitly"

No.	Observation	Conversant	Verbatim
1	VO Typing 'explicitly' in EE	Researcher	What were you thinking when you looked for 'explicitly' in EE?
2		JWK	I wanted to find out the definition.
3		Researcher	What were you looking at?
4		JWK	I saw the definition. I saw the first line and found out the definition.
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs, raising doubts about whether the fitful negotiations are making even limited progress.			
RS: The clear rejection by North Korean radio on Saturday the formula which the U.S. has basically tried to finish negotiations about the North Korea's nuclear program is increasing skepticism about whether the intermittent talks has advanced in a small portion.			

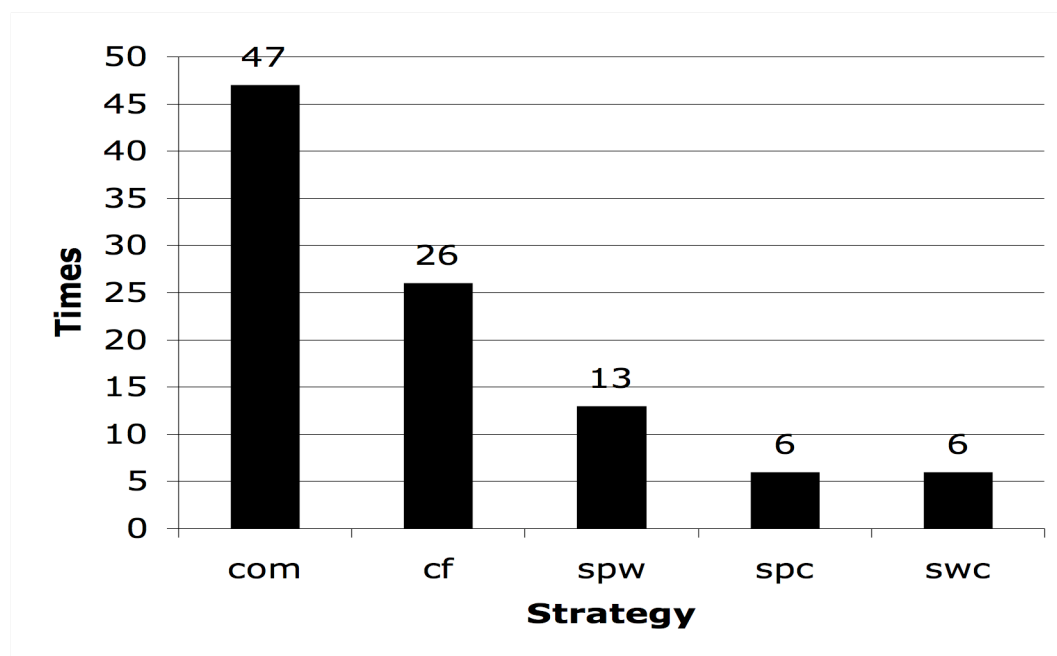
When looking at the totals of the searches performed for the purpose of samples, verbs, definitions, etc., it is easy to see that the preposition was the most common information investigated (52 times, 14.2%). Interestingly, when the subjects searched for prepositions, the search words they used were not prepositions, they were the verbs that they believed the prepositions followed. For this reason, when they did further searches,

the combination was verb + preposition. Oftentimes, subjects looked at the collocation frequencies to see how many times the collocations they were looking for occurred. Second in line were verbs (51 times, 13.9%), with nouns coming in third place with 42 searches. When subjects used the Korean-English dictionary, the information they looked for most is definitions. When thesauri were used, most of the time the subjects were searching for synonyms of verbs to change the verbs from the original text because their task was to paraphrase the newspaper article.

4.3. What strategies do learners use when they consult a corpus?

Research data for this question were obtained from the screen recording and during the stimulated recall session. The screen recording showed, in part, what strategies the subjects chose to use for better results. Strategic behaviors that could not be observed on the screen recording were elicited during the stimulated recall. Research data were obtained in response to interview questions such as "What were you thinking when you changed the search word?" Figure 4.4 shows that the most common strategy is using a combination of reference tools (47 times, 48.0%). This number does not include cases when the same tool was used multiple times consecutively. The second most common strategy is referring to the collocation frequency data that *MonoConc* provided (26 times, 26.5%). Other strategies include: using a more specific search term when the initial search results in too many or irrelevant examples; using special characters in order to see examples with all possible forms of the search term or with two search words that are not next to each other; changing the search word when the subjects thought the results were not relevant to the context or their initial reason for the search. This section shows how the subjects implemented these strategies when using *MonoConc*.

Figure 4.4. Strategies for using a concordancing program



com: combination of tools, cf: using collocation frequency list, spw: using more specific search term, spc: using special character, swc: search word change

Subjects implemented strategies while using their reference tools by combining tools. Sequences of tool combinations include dictionary-*MonoConc*, thesaurus-*MonoConc*, and vice versa, in addition to *MonoConc-MonoConc*. Another strategy was changing search terms in *MonoConc*, such as a broad search to specific search or the reverse. This means that subjects could look for a wide range of examples and narrow the examples down to only a select few, or vice versa. In addition to changing search words to a broad or narrow range of examples, subjects also used symbols such as the asterisk (*) and the “at” symbol (@).

In Table 4.17, the participant first typed ‘reject’ in the thesaurus. The results showed ‘decline,’ so she copied it into *MonoConc*. In *MonoConc*, she checked to see if it was used in a political context and if it was a noun or verb (see line 9). She also confirmed the kinds of words that came before and after ‘decline.’ She wanted to write ‘decline + noun.’

Table 4.17. Subject HJS using thesaurus with search term "reject" and *MonoConc* with search term "decline"

No.	Observation	Conversant	Verbatim
1	VO Reading the article. VO Typing 'reject' in T	Researcher	Reject.
2		HJS	I was going to change it to a different word. I wanted to change 'reject' to a different word.
3	VO Showing results in T	Researcher	What were you looking at?
4	AC Pointing at results in T	HJS	I saw 'decline.' I copied 'decline.'
5	VO Copying 'decline' and pasting 'decline' into M	HJS	I was wondering if I could use 'decline.'
6		Researcher	Why did you look for 'decline'?
7		HJS	To see if I could use it in this context, and to see the part of speech.
8		Researcher	What kind of context were you looking for?
9		HJS	For example, I wanted to see whether it was used as a noun or verb. And I wanted to see if 'decline' could be used in a political context.
10	VO Looking at collocation frequency (CF)	Researcher	Why did you look for CF?
11		HJS	To see what kinds of words come before and after it. I thought that it would be better to use a sentence structure that was used frequently by other people. I was going to write 'decline' and a noun. I was looking to see whether a noun was frequently used after 'decline.'
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: An announcement of North Korean radio declined the offer of the United States that North Korea should stop all its nuclear programs.			

Immediately following the example in Table 4.17, this same participant (see Table 4.18) typed 'declin*' in *MonoConc*. She was looking for 'Decline health' or something similar ('decline + noun') and found it in *MonoConc*. The correct choice

would have been ‘declining health,’ but by using the character (*), she was able to find that sequence in the corpus. Had she typed in ‘decline health,’ she might not have found the information she was searching for.

Table 4.18. Subject HJS using *MonoConc* with search term "declin"

No.	Observation	Conversant	Verbatim
1	VO Typing 'declin*' in M	HJS	I looked for it again.
2		Researcher	What were you looking at?
3		HJS	'Decline health' and something similar, 'decline' followed by a noun.
4	VO Looking at CF	HJS	In order to save time, I looked at this.
5		Researcher	What did you decide?
6		HJS	As I thought, 'decline' should be followed by a noun.
7		Researcher	Where did you get that idea?
8		HJS	It was the second highest frequent form, so I thought I could use it. Generally, 'decline' can be followed by a noun, so I was just checking. I didn't have to see any more because it was the second highest frequent form.
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: An announcement of North Korean radio declined the offer of the United States that North Korea should stop all its nuclear programs.			

The episode recorded in Table 4.19 shows the participant entering ‘a visit’ into *MonoConc*. She saw ‘a visit to North Korea.’ and realized that she usually used ‘visit’ as a verb, as in the simple past tense form ‘visited.’ She decided to change her *MonoConc* inquiry to ‘a visit’ in order to only show nouns. She saw various examples and saw a similar structure to that for which she was searching. Line 10 indicates that she was going to write ‘visit to Korea.’ However, her final decision was to write “. . . the visit of

China's foreign minister . . . ” which was also an acceptable expression. Her strategy had been to use a specific search word.

Table 4.19. Subject JDO using *MonoConc* with search term "visit"

No.	Observation	Conversant	Verbatim
1	VO Typing 'a visit' in M	Researcher	Why were you looking for 'a visit' in M?
2		JDO	I saw 'a visit to North Korea.' I usually used 'visit' as a verb like 'visited' and not as a noun. The reason I revised a lot is that the sentence becomes too long if I write 'just after China's Foreign Minister visited North Korea.' So, I thought I should change the verb, 'visit,' to a noun, 'a visit.' I was wondering how 'a visit' was used.
3		Researcher	Why did you add 'a'?
4	VO Canceling the search while M is searching	JDO	To make sure that what I'm looking for is a noun. If I just look for 'visit,' it'll also show verbs.
5		Researcher	What were you looking for?
6		JDO	A visit to China,' 'a visit to Notre Dame,' and 'a visit to . . . ' They are 'a visit to' something. When using 'visit' as a verb, it's 'visit China.' When 'visit' is a noun, it becomes 'visit to.' Here, it says, 'visited Notre Dame.'
7	AC Reading the article	JDO	Here it is. It's the Foreign Minister who is going to visit. The content of the examples are similar to this sentence.
8	AC Pointing at the results in M	JDO	Like in this example, 'A visit by Taiwan President Lee,' I was wondering if I could write 'a visit by the Foreign Minister.' I was looking for sentences with a similar structure, and I found this sentence.
9		Researcher	What did you learn from this sentence?
10		JDO	I could write 'visit to Korea.'
OS: The statement carried by Radio Pyongyang and monitored by news agencies in South Korea came just after a visit to North Korea by China's foreign minister, Li Zhaoxing . . .			
RS: The release of public statement which presented by Radio Pyongyang and monitored by news agencies in South Korea was done just after the visit of China's foreign minister Li Zhaoxing to North Korea . . .			

The fourth strategy used was a search word change. At first, this subject (see Table 4.20) typed ‘carry’ in the thesaurus. She wanted to use a different word, but could not find anything appropriate. Finally, she decided to use ‘announce,’ but later changed it to ‘monitor’ and still could not find any words she liked.

Table 4.20. Subject JWK using thesaurus with search terms "carry" and “monitory”

No.	Observation	Conversant	Verbatim
1	VO Typing 'carry' in T	Researcher	What were you thinking when you looked for 'carry' in T?
2		JWK	I wanted to use a different word instead of 'carry.' But I didn't find any appropriate words that I liked to use.
3		Researcher	So what did you do?
4		JWK	I used a different word. I used 'announce.' I think I changed this word later. I knew the meaning of 'carry,' but it was difficult to change it to a different word. So, I was looking for different words but didn't find any.
5	VO Typing 'monitory' in T	JWK	That's why I looked for 'monitor' also, but I still didn't find anything I wanted.
6		JWK	I knew 'monitory' as surveillance (in Korean), but I couldn't find any words I liked.
OS: The statement carried by Radio Pyongyang and monitored by news agencies in South Korea came just after a visit to North Korea by China's foreign minister, Li Zhaoxing . . .			
RS: This statement given by Radio Pyongyang and watched over by news agencies in South Korea announced just after coming to North Korea by Li Zhaoxing, China's foreign minister, and shortly before April . . .			

The most common macro strategy used was combining tools (41 times). The second most used was using a specific search word (13 times). Search word changes and special characters were each used on six occasions. It appears, then, that the subjects had the most success when they used more than one tool for a search. More often than not, one search would lead to multiple searches, until the subjects found what they considered to be the most appropriate option.

In Table 4.21, the subject relied on *MonoConc* to make his final decision about what should be used with ‘insist.’ He wanted to see if ‘insist that’ was possible or not in order to paraphrase. When he did not see that example, he went with the original phrase ‘insist on,’ which he already knew.

Table 4.21. Subject KWA using *MonoConc* with search term "insist"

No.	Observation	Conversant	Verbatim
1	VO Typing 'insist' in M	Researcher	You came back to M and typed in 'insist.'
2		KWA	I already know 'insist on,' but I wanted to see if it was right or wrong. I was also checking to see whether an expression like 'insist that' was possible or not.
3		Researcher	Why were you looking for that information?
4		KWA	For paraphrasing, I wanted to use a that-clause that says the North Korean government rejects the US government's position on the nuclear programs.
5		Researcher	You wrote 'insist on.' Why did you decide to write this?
6	VO Mouse is pointing at 'insist on pressing' in W	KWA	At the end I decided not to write it. I saw 'on' in M, so I used that instead.
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs, raising doubts . . .			
RS: The government of North Korea blamed for the policy the United States insists on giving up their nuclear programs through their radio broadcasting.			

The same subject relied on *MonoConc* to make another decision. He decided not to write ‘press’ because he did not find an example of what he was looking for. (See Table 4.22.)

Table 4.22. Subject KWA *MonoConc* with search term "press"

No.	Observation	Conversant	Verbatim
1	AC Pointing at 'constantly pressing him to find' in M	Researcher	Oh, you didn't see this?
2		KWA	I didn't see that. I don't think I used 'press' at the end.
3		Researcher	Why did you not use it?
4		KWA	I didn't see that sentence. Usually I write like that, but I wanted to use correct grammar this time. I gave up because I didn't see that. I only saw a lot of examples with 'pressure.'
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: The government of North Korea blamed for the policy the United States insists on giving up their nuclear programs through their radio broadcasting.			

This third example (see Table 4.23) is shown from the same subject because he was again influenced by the examples that appeared in *MonoConc*. He chose not to use 'criticize' because many samples of 'say + that-clause' were given. His final decision was to use "... said that ..." which was correct, while other parts of the phrase sounded awkward.

In Table 4.24, the same participant as shown in Tables 4.21 through 4.23 again relied on the results of *MonoConc* to decide if what he wanted to use was correct or not. He wanted to use 'make an agreement,' but could not find any results (see line 2) and believed his phrase to be incorrect.

Another subject gives us a look into her use of characters along with words and phrases in *MonoConc*. (See Table 4.25.) At first, she typed 'announc* that' in *MonoConc* (see Table 4.25). She wanted to know if she could use 'announce to,' but 'announce that' was used more often. She typed 'announc* to' later (line 5), and made comparisons between the two. She saw that 'announce + that-clause' was used more

often and it seemed to be better. She decided to write what to announce instead of writing about who was going to be doing the announcing.

Table 4.23. Subject KWA using *MonoConc* with search term "reproach"

No.	Observation	Conversant	Verbatim
1	VO M is displaying results.	Researcher	What else are you looking at?
2		KWA	I was waiting for more examples to show up, but I never saw examples with that-clauses.
3		Researcher	Let's see what you did.
4		KWA	I didn't use 'criticize.' Instead, I used 'say' because 'say + that-clause' appeared a lot. So, I just decided to use it.

OS: "The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime," it said, using the initials of North Korea's official name, the Democratic People's Republic of Korea.

RS: A North Korea's official said that the present policy of the U.S. just makes the Korean peninsula in danger, which can occur a nuclear war.

Table 4.24. Subject KWA using *MonoConc* with search term "make an agreement"

No.	Observation	Conversant	Verbatim
1	VO Typing 'make an agreement' in M	Researcher	What were you thinking when you looked for 'make an agreement' in M?
2		KWA	I think I have heard of it. I didn't want to use 'agree' for 'dong-i' (agree). So, I was wondering whether 'make an agreement' is correct. There weren't any results, so I thought it was wrong.

OS: The administration has also said it will not provide aid or other benefits to North Korea before it scraps all its nuclear programs and allows rigorous inspections.

RS: The Bush administration responded to it by saying they will not reach an agreement and support North Korea without any clear and rigorous inspections.

Table 4.25. Subject JYP using *MonoConc* with search term "announc*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'announc* that' in M	Researcher	What were you thinking when you looked for 'announce' in M?
2		JYP	I wanted to know if I could use 'announce to.' At first, I was going to write 'announce that.' I wanted to see if I could use 'announce to.' It seemed that 'announce that' was used more.
3		Researcher	What were you looking at?
4		JYP	I was looking at how many times 'announce that' was used and compared the frequency with that of 'to.'
5	VO Typing 'announc* to' in M	Researcher	This time, you were looking for 'announce to.'
6		JYP	There were more cases that a noun came after it. I was wondering whether to use 'announce to' or an 'announce that-clause.'
7		Researcher	What did you pay attention to?
8		JYP	If a noun followed more frequently, and if the context of the examples were similar to the context of the sentence I was going to write it.
9		Researcher	What were you comparing between 'announce that' and 'announce to'?
10		JYP	'Announce to' is usually followed by a person or a noun, but not by a verb. So, I decided to use a that-clause.
11		Researcher	You already wrote 'announce that' in W.
12		JYP	I was just checking.
13	VO Looking at CF of 'announ* to' in M	Researcher	Why were you looking at the frequency?
14		JYP	To see if a verb or noun came after 'to' more frequently. There weren't many cases that 'announce that' was followed by a verb, so I used a that-clause. 'Announce + that-clause' was used more frequently, and it sounded better. Also, I didn't have to say to whom to announce. Instead, I was going to write what to announce.

OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks . . .

RS: North Korean radio explicitly announced that they decided not to accept the formula the United States has suggested as its basic position in talks . . .

Table 4.26. Subject JYP using *MonoConc* with search term "blame*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'blame* that' in M	Researcher	Why were you looking for 'blame* that' in M?
2		JYP	Whether 'blame' could be followed by 'for' like 'accuse,' and if 'blame + that-clause' was used. I was also looking for which expression was used more frequently.
3		Researcher	What did you find?
4	VO Showing 6 results in M. AC Pointing at 'He blames that same negativity . . . ' in M	JYP	This 'that' is not a part of a that-clause. It's a part of 'that same.' So, I thought 'blame + that-clause' was not used. I looked for 'blame + a person + for.'
5		Researcher	Did you not get what you want?
6		JYP	I was just checking. I was looking to see how to use 'blame that' and 'blame for.' But 'blame that' was not used.
7	VO Typing 'blam* for' in M	JYP	There were a lot of samples with 'blame for,' so I thought about using it.
8		Researcher	What were you looking at?
9		JYP	Reasons for blaming come after.
10	VO Typing 'blam*@for' in M	JYP	I was also looking for 'blame + a person + for,' and I found some examples. Previously, I looked for 'blame for.' I was wondering if I could put a person in between. According to the examples, it's possible.
11	VO Typing 'blamed the US hawks and followers for the danger' in W	JYP	So, I thought I could write 'blamed the US hawks and followers for.'
OS: The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime . . .			
RS: North Korea blamed the U.S. hawks and followers for the danger of the Korean Peninsular.			

In the episode shown in Table 4.26, the subject typed 'blame* that' in *MonoConc*. She wanted to see if 'blame for' could be used as 'accuse,' and if 'blame + that-clause' was used. The 'that' that she found was not part of a that-clause. After this, she typed 'blam* for' in *MonoConc* (see line 10) to see if reasons for blaming could follow the verb. She also typed 'blam*@for' in *MonoConc* to see if a person could be put within

the phrase. She found out this was possible, and so she decided to write ‘blamed the US hawks and followers for.’ As students gained more experience while using *MonoConc*, they became more confident in employing strategies that would lead them to make the best decisions for their writing. As experiments go, sometimes the learners were successful, and sometimes they were not.

In summary, when subjects use the combination of dictionary and *MonoConc*, they find out the meanings or synonyms in the dictionary and then used *MonoConc* in order to confirm that what they have written is correct. They needed to see if any of the examples had the same structure. In this case, *MonoConc* serves as a quick reference for confirmation. When using the combination of *MonoConc-MonoConc*, it is because they need to either narrow down the searches or obtain a broader range of examples. When there are too many examples, they are not always able to find what they want to see. When this happens, they narrow the search to find what they need. Conversely, if the search is narrowed too much or if the search terms are too specific, they will obtain either too few or no results. This can be remedied by broadening the search terms and or employing a special character (*). Sometimes when no example is given, the subjects perceive that what they write is not correct.

4.4. How does a corpus complement the use of dictionaries?

Research data for this question were obtained during the stimulated recall session in response to interview questions such as "Why did you use the same search word in *MonoConc* after using the English-English dictionary?" Anticipated possible ways of complementing dictionary use included: needing to know collocations other than those shown in the dictionary; needing to see use of the search word in a specific context which is not found in dictionaries; needing to get examples of using a phrase in a sentence

quickly; needing to see more examples in order to understand the meaning of a word and grammatical structure in context.

In many ways, corpora use can help second language writers by providing features that dictionaries and thesauri do not have, such as numerous real-life examples that writers can use to draw conclusions about phrases or sentence structures. By using corpora, writers can also find the context in which search words are used. The concordancing program in this study also provides easy access to the list of collocation frequencies so that users can find out what the surrounding words are. Dictionaries provide examples and show some phrases that can be used with the search word. However, they provide only a limited number of examples and collocations. This section focuses on how *MonoConc* was used in conjunction with and as a complement to the other tools.

4.4.1. Combination of thesaurus and *MonoConc*

In Table 4.27, as shown above, the subject uses a combination of thesaurus and *MonoConc*. The thesaurus is used first, followed by *MonoConc*. In this example, the participant was initially looking for another word for ‘reject.’ She looked it up in the thesaurus first and saw the word ‘decline.’ Then, she typed ‘decline’ in *MonoConc* in order to find out how to use it in a sentence and in a political context (see line 9). She also looked at the collocation frequency (see line 19) to find out the part of speech of words that come after ‘decline.’ She found out that the article ‘the’ could be frequently used as a word that comes after ‘decline.’ After looking at some more examples of this combination, she typed ‘declin*’ in *MonoConc*. She realized that ‘decline’ should be followed by a noun. Finally, she typed ‘declined the United States’ as her final answer.

MonoConc complements thesaurus use by providing quick access to collocation frequency, which a thesaurus alone cannot provide. Then, *MonoConc* helps the subject decide on the expression to be used by providing many examples.

Table 4.27. Subject HJS using thesaurus with search term “reject” and *MonoConc* with search term "decline"

No.	Observation	Conversant	Verbatim
1	VO Typing 'reject' in T	Researcher	Reject.
2		HJS	I was going to change it to a different word. I wanted to change 'reject' to a different word.
3	VO Showing results in T	Researcher	What were you looking at?
4	AC Pointing at results in T	HJS	I saw 'decline.' I copied 'decline.'
5	VO Copying 'decline' and pasting 'decline' into M	HJS	I was wondering if I could use 'decline.'
6		Researcher	Why did you look for 'decline'?
7		HJS	To see if I could use it in this context, and to see the part of speech.
8		Researcher	What kind of context were you looking for?
9		HJS	For example, I wanted to see whether it was used as a noun or verb. And I wanted to see if 'decline' could be used in a political context.
10	VO Looking at CF	Researcher	Why did you look for CF?
11		HJS	To see what kinds of words come before and after it. I thought that it would be better to use a sentence structure that was used frequently by other people. I was going to write 'decline' and a noun. I was looking to see whether a noun was frequently used after 'decline.'
12	VO 'in'- most frequent word 1-right in CF	HJS	I saw 'in.'
13	AC Pointing at 'the'- most frequent word 2-right	HJS	I can use a noun because 'the' comes as a second word to the right. So, I thought I could use this structure.
14		Researcher	What were you looking at in the results?
15	AC Pointing at 'decline the' in results in M	HJS	I saw 'decline the.'
16	VO Typing 'declin*' in M	HJS	I looked for it again.
17		Researcher	What were you looking at?

Table 4.27. Continued

18		HJS	'Decline health' and something similar, 'decline' followed by a noun.
19	VO Looking at CF	HJS	In order to save time, I looked at this.
20		Researcher	What did you decide?
21		HJS	As I thought, 'decline' should be followed by a noun.
22		Researcher	Where did you get that idea?
23		HJS	It was the second highest frequent form, so I thought I could use it. Generally, 'decline' can be followed by a noun, so I was just checking. I didn't have to see any more because it was the second highest frequent form.
24	VO Typing 'declined the United States' in W		

OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .

RS: An announcement of North Korean radio declined the offer of the United States that North Korea should stop all its nuclear programs.

In the following example (see Table 4.28), the subject used the thesaurus to look for synonyms of 'inconclusive.' He found 'ambiguous' and decided to use it. He then used *MonoConc* to find the appropriate preposition to be used after 'ambiguity.' After he used the thesaurus, the subject tried to find the correct preposition by reading the results. He could not find any examples with 'ambiguity to' (see line 5) or 'ambiguity for' (see line 7) and discovered that neither was correct. Instead, he found 'ambiguity of.' Then, he decided to use a specific search phrase in order to find a specific example containing 'ambiguity to.' He also searched for 'ambiguity for' in order to clarify that it was not correct. He then wrote 'ambiguity of' in his paraphrase. However, the final result was that only 'ambiguity' was used. In this example, *MonoConc* was used to clarify the subject's impression. In a case such as this, using only a thesaurus could drastically limit the examples offered in a search. It would therefore be very difficult, if not impossible,

to answer questions such as these. On the other hand, *MonoConc* offers a variety of examples, giving the user ample opportunities for making a well-informed decision. In this case, *MonoConc* helped the user to decide not to write the phrase that he first intended to use. Typically, *MonoConc* is useful because it provides numerous examples for the user. This time, *MonoConc* did not provide very many examples and thus helped the user to make a decision based on lack of evidence.

Table 4.28. Subject MKS using thesaurus with search term “inconclusive” and *MonoConc* with search term "ambiguity"

No.	Observation	Conversant	Verbatim
1	VO Typing 'inconclusive' in T	Researcher	So, you were looking for 'inconclusive' in T.
2		MKS	I saw 'ambiguous,' so I thought I could use it.
3	VO Typing 'ambiguity' in M	Researcher	You looked for 'ambiguity' while you were writing 'ambiguity' in W.
4		MKS	Because I wasn't sure whether 'ambiguity to' or 'ambiguity for' was right. I used M to look for a correct preposition, and I found out both were wrong. There wasn't 'ambiguity to' or 'ambiguity for.' But, there was 'ambiguity of.'
5	VO Typing 'ambiguity to' in M	Researcher	You were looking for 'ambiguity to' in M.
6		MKS	I was still wondering if 'to' was correct.
7	VO Typing 'ambiguity for' in M	Researcher	So, you were looking for 'ambiguity for' this time. But there were no results. What did you do?
8	VO Typing 'of' in W	MKS	I used 'of' because I saw that 'of' was used.
9		Researcher	You didn't use 'ambiguity of.' Instead you wrote 'ambiguity' and a period.
10		MKS	I didn't decide whether to write 'of the effort' after 'ambiguity.' I thought I would decide on it later and moved on. I don't remember if I changed it.
OS: But in the latest statement, it appears to be setting the stage for another inconclusive effort.			
RS: Yet in the latest statement last Saturday reveals an ambiguity.			

4.4.2. Combination of English-Korean dictionary and

MonoConc

The next most frequent combination of reference tools used was the English-Korean dictionary (EK) and *MonoConc*. Here are some examples of how *MonoConc* complements dictionary use. In the following example, the subject used EK to see what kinds of words follow the word ‘reject.’ She found that a noun could follow the verb (see line 6) and used *MonoConc* to clarify whether nouns or infinitives follow the verb ‘reject’ (see line 7). She decided to use a noun after looking at examples in *MonoConc*. She deleted ‘to use’ and typed ‘the use’ after ‘reject.’ Although ‘the use’ is not grammatically correct, she discovered that a noun should come after ‘reject’ by reading the examples. See details in Table 4.29.

Table 4.29. Subject JYP using English-Korean dictionary with searcher term “reject” and *MonoConc* with search term "reject*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'reject' in EK	Researcher	What were you looking at?
2		JYP	I saw that 'reject' was followed by a noun.
3	VO Typing 'reject*' in M	Researcher	You were looking for 'reject' in M again.
4		JYP	I was wondering what kind of words could come as an object after 'reject.'
5		Researcher	What were you looking at?
6		JYP	I found out that nouns came after it, not verbs. So, I used a noun.
7		JYP	I found out that nouns came after it, not verbs. So, I used a noun.
8	VO Deleting 'to use' and Typing 'the use' in W	JYP	So, I deleted 'to' and changed to 'rejected the use.'
OS: The statement rejected the American formula point by point.			
RS: However, North Korea rejected the use of that wording.			

In the next set of data (see Table 4.30), the subject used *MonoConc* after using the English-Korean dictionary (EK) to find appropriate sentence or phrase structures. First, she used EK to find out whether a that-clause or ‘for + person’ should be used after the verb ‘trust.’ She found ‘trust him for (or to do) that’ in EK (see line 3). She was going to write ‘trust + object + for,’ but she changed her mind and used a that-clause instead. She utilized *MonoConc* to see if a that-clause should be used, but she did not find it. She found that ‘to + verb’ form was used in the examples. She then looked at the collocation frequency of ‘trust’ and found out that the word ‘that’ was most frequently used after ‘trust’ in the corpus (see line 7). She went back to look at the data in *MonoConc* to see if the that-clause was really used in the sample sentences. She found some examples of that-clauses in *MonoConc*, so she decided to use it in her sentence. She changed the structure ‘trust + object + to’ to ‘trust + that-clause’ (see line 11). This example illustrates that EK showed the possibilities of both sentence structures, but was unclear which one was the best choice. *MonoConc* helped her decide which one to use for her situation by collocation and examples.

The following example in Table 4.31 shows the use of *MonoConc* as a tool that confirms what the subject first found in the English-Korean dictionary. He searched for ‘boon’ in EK first and saw that it was related to ‘benefit.’ Immediately following this, he typed ‘boon’ in *MonoConc* in conjunction with using it in his paraphrase. The results in *MonoConc* showed ‘economic boon to the US.’ He saw that ‘boon + to’ was a pattern found in both *MonoConc* and EK (see line 6). *MonoConc* confirmed the solution to this subject because he felt that he could not trust Korean-English dictionaries because they tend to give examples that are not authentic and he likes to see authentic examples found in *MonoConc*.

Table 4.30. Subject MAP using English-Korean dictionary with search term “trust”
MonoConc with search term "trust"

No.	Observation	Conversant	Verbatim
1	VO Typing 'trust' in EK	Researcher	You were looking for 'trust' in EK.
2		MAP	I was looking to see if a that-clause or 'trust for + person' was correct.
3	AC Pointing at 'trust him for [or to do] that' in EK	MAP	Like this here, 'trust + object + to.'
4	VO Typing 'trust' in M	Researcher	You were looking for 'trust' in M.
5		MAP	I was going to write 'object + for,' but then the sentence became too long. So, I was checking if a that-clause could be used.
6	VO Typing 'trust America to' in W	Researcher	How did you decide to use 'to'?
7		MAP	I didn't find a that-clause, so I was going to use 'to.' I then checked CF and found 'that.'
8	VO Looking at CF. CF showing 'that' occurred 9 times as 1-right word	MAP	So, I thought a that-clause was possible.
9	VO Scrolling down the results in M	MAP	I was looking for a that-clause.
10	VO M showing examples with that-clause		
11	VO Changing 'trust America' to 'trust that . . . ' in W		
OS: . . . raising doubts about whether the fitful negotiations are making even limited progress.			
RS: North Korea doesn't trust whether the talk is really working.			

Table 4.31. Subject MKS using English-Korean dictionary with searcher term “boon” and *MonoConc* with search term "boon"

No.	Observation	Conversant	Verbatim
1	VO Typing 'boon' in EK	MKS	I found out that 'boon' was used. It's related to 'benefit.'
2	VO Typing 'boon' in M	Researcher	You looked for 'boon' in M when you were writing 'boon' in W.
3		MKS	I just wanted to see if 'boon' was really used, and how it was used.
4	AC Reading the results in M	MKS	I saw 'economic boon to the US.' So, I thought that 'boon' was used. That's it.
5		Researcher	How did you know 'boon + to'?
6		MKS	I found it in M and EK. The biggest reason for using M is because I don't trust Korean-English dictionaries.
7		Researcher	What do you not trust about them?
8		MKS	They provide definitions okay, but the examples seem to be Konglish (Korean + English). I knew that some expressions are Japanese style grammar. I don't know how dictionaries are made these days, but I use M to see how Americans use expressions.
OS: The administration has also said it will not provide aid or other benefits to North Korea before it scraps all its nuclear programs and allows rigorous inspections.			
RS: The administration has said that it will not offer any aids or other boons to North Korea before it closes all its nuclear programs and consents punctual inspections.			

The next combination of reference tools used was the Korean-English dictionary (KE) and *MonoConc*. Here are some examples of how *MonoConc* complements dictionary use. In the following example, *MonoConc* is used in conjunction with KE for understanding sentence structures. The subject first used KE to look for English equivalents of a Korean word (chujanghada). She found the English translation ‘assert.’ She decided to look for ‘assert’ in *MonoConc* in order to see how it was used in a sentence. She found that ‘assert’ was followed by a that-clause, unlike ‘insist,’ which she explained, was followed by ‘on.’ See Table 4.32.

Table 4.32. Subject JDO using Korean-English dictionary with search term “chujanghada” and *MonoConc* with search term "assert"

No.	Observation	Conversant	Verbatim
1	VO Typing 'chujanghada' (claim in English) in KE	Researcher	Why were you looking for 'chujanghada' (claim in English) in KE?
2		JDO	I didn't want to use 'insist' over and over. There is 'suggest' for the same meaning, but 'insist' has a stronger meaning than 'suggest.' I also thought about using 'propose.'
3		JDO	Then, I thought that I should look for 'chujanghada,' so I could find related words.
4		Researcher	What were you looking at?
5		JDO	I didn't like any of the words I found.
6	AC Pointing at results in KE		'Affirm' and 'assert' were not what I wanted.
7	VO Typing 'assert' in W	Researcher	Then why did you write 'assert' in W?
8		JDO	I didn't like to use 'insist.' I thought 'assert' was a little better.
9	VO Typing 'assert' in M	JDO	I wanted to know how 'assert' was used. 'Insist' is always followed by 'on,' so it's 'insist on something.' But it's 'assert something.' I found out that there wasn't any word like 'on' that followed 'assert.'
10		Researcher	What were you looking at?
11		JDO	'Assert' is followed by a that-clause. An object follows 'assert,' not like 'insist on something.'
OS: It put North Korea on record as saying that it could not accept the main goals President Bush and his negotiators have insisted on in the first two rounds of talks.			
RS: North Korea established its position that it won't accept the primary goals that President Bush and his followers have asserted in the first two rounds of talks.			

In the next example (see Table 4.33), *MonoConc* was used in a similar way in conjunction with KE. First, the student used KE to find the translation for ‘taeung’ (reaction) in English. The context he was looking for dealt with how the US reacted to or responded to what North Korea announced. He wanted to use ‘react,’ but decided to look for other words. Some definitions he saw included ‘confrontation and opposition,’ but he

Table 4.33. Subject KWA using Korean-English dictionary with search term “tae-ung” and *MonoConc* with search term “respond*”

No.	Observation	Conversant	Verbatim
1	VO Typing 'tae-ung' (reaction) in KE	Researcher	You are looking for 'tae-ung' (reaction) in KE.
2		KWA	The US reacted or responded to what North Korea announced. This kind of expression is what I wanted to use. I was going to use 'react,' but I didn't like it. So, I was looking for other words.
3	AC Reading definitions in KE	KWA	The first thing I saw was 'confrontation and opposition,' but I didn't like it. I also saw those expressions down below, but I didn't find what I wanted.
4	VO Typing 'ung-dap' in KE. AC Reading definitions in KE		That is why I was looking for 'ung-dap' (response). I saw 'response' and used it.
5	VO Typing 'respond*' in M	Researcher	You are looking for 'respond' in M.
6		KWA	I was looking at the sentence structure. For example, I wasn't sure if 'respond to' is correct, or 'respond' is correct. I wanted to borrow sentence structures and use them.
7		Researcher	What structures are you looking for?
8		KWA	Not any particular structures. I found that 'respond' and 'to' were used together. I might have found a sentence structure that I copied into my writing.
9		Researcher	Where is the sentence?
10	VO Looking at CF	Researcher	You are also looking at the frequency.
11		KWA	To make sure 'to' can be used. I was also looking for 'that.'
12	AC Pointing at 'initially responded to the furor by saying' in results in M		Here, 'responded to' is used as a reaction to what was said. So, I copied 'responded to' and 'by saying' in my writing.
13		KWA	As a result, my sentence says that the US responded to North Korea by saying something. So I just copied it into my writing because I didn't want to use 'announced.' I wanted to replace it.
14	VO Typing 'responded to it by saying' in W	KWA	So, I saw this sentence and copied it.

Table 4.33. Continued

OS: The administration has also said it will not provide aid or other benefits to North Korea before it scraps all its nuclear programs and allows rigorous inspections.

RS: The Bush administration responded to it by saying they will not reach an agreement and support North Korea without any clear and rigorous inspections.

did not like those words either. The subject decided to type another word into KE (ungdap) (line 4). ‘Ungdap’ means respond, so he used this result and typed ‘respond*’ in *MonoConc*. This time he was looking for sentence structure because he was not sure if ‘respond to’ or ‘respond’ is correct. He found that ‘respond to’ is correct and checked out the collocation frequency to confirm that ‘to’ could be used and to also look for ‘that.’ He found an example that stated ‘initially responded to the furor by saying,’ so he used ‘responded to’ and ‘by saying’ in his writing (line 13). Because of these findings, the final version includes “responded to it by saying.”

In the example shown in Table 4.34, the subject used KE for the purpose of looking for a slang term. She typed the word ‘hasu’ in KE and found the word ‘underdog,’ but did not find the translation helpful. Next, he decided to look for ‘tolmani’ in KE and saw the word ‘underling.’ This subject said that she did not trust Korean-English dictionaries, so she typed ‘underling’ in M and found that it was a disrespectful term (see line 10) just as the Korean words she searched for in KE, so she used ‘underling’ in her paraphrase.

In the search shown in Table 4.35, the same subject performed another search in KE with the word ‘nagada,’ which means ‘step forward.’ The subject was trying to replace the phrase ‘push forward’ in the article. She also typed ‘step forward’ in M to see how it is used in sentences and found that it is usually used as a noun. She saw the example ‘to step forward on this issue’ and saw that it could mean stepping forward physically or a situation that is progressing (see line 5). She wanted to find another

example with a similar structure and context. She typed ‘step forward on third table’ in her paraphrase and said that she knew to use ‘on’ because the phrase is about stepping on or moving onto something. She pointed at the same example again ‘step forward on this issue’ in M to show that this was the phrase that gave her the information she needed (see line 8).

Table 4.34. Subject MKS using Korean-English dictionary with search term “hasu” and *MonoConc* with search term "underling"

No.	Observation	Conversant	Verbatim
1	VO Typing 'hasu' ('underdog' in English) in KE	Researcher	You were looking for 'hasu' ('underdog' in English) in KE.
2		MKS	I was going to look for a slang term for that word.
3	VO Typing 'tolmani' ('underdog' in English) in KE. AC Reading the article	MKS	I didn't get much out of it when I looked for 'hasu' ('under dog' in English), so I looked for 'tolmani.'
4		MKS	In the article, it says, 'the US war hawks and their followers.' I was looking for disrespectful expressions, so I found 'underling.'
5	VO Typing 'underling' in M	Researcher	You were looking for 'underling' in M.
6		MKS	I wanted to see how it was used in real life because I didn't agree with Korean-English dictionaries.
7		Researcher	What did you learn here?
8	VO Clicking on an example and reading the full context	MKS	I clicked on one example because it was cut off.
9		Researcher	What were you looking for?
10	VO Typing 'their underlings is driving' in W	MKS	I found out that 'underling' was used as a disrespectful term like the Korean words I looked for in KE. So, I used 'underling.'
OS: The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime . . .			
RS: The headlong moves of the US warmongers and their underlings are throwing the Korean peninsula into a dangerous vortex in which a nuclear war against the DPRK may break there anytime.			

Table 4.35. Subject MKS using Korean –English dictionary with search term “nagada” and *MonoConc* with search term “step forward”

No.	Observation	Conversant	Verbatim
1	VO Typing 'nagada' ('step forward' in English) in KE	Researcher	Why were you looking for 'nagada' ('step forward' in English) in KE?
2	AC Pointing at the article	MKS	Here it says, 'push forward.' I knew it was 'proceed' or 'advance.' But it's a phrase, so I was looking for a phrase to replace this phrase.
3	VO Typing 'step forward' in M	Researcher	Why were you looking for 'step forward' in M?
4		MKS	I wanted to see how it's used in sentences. Usually, it was used as a noun.
5	AC Pointing at 'to step forward on this issue' in M	MKS	Here, I saw 'to step forward on this issue.' So, I thought I could write like this. I copied it. It could be stepping forward physically, or a situation is progressing. So, I wanted to see examples to find a sentence with similar structure and context.
6	VO Typing 'step forward on third table' in W	Researcher	How did you know to use 'on'?
7		MKS	Because it's about stepping on or moving onto something.
8	AC Pointing at 'step forward on this issue' in M	MKS	I saw this sentence with 'on.'
OS: Mr. Li said this week that North Korea was ready to "push forward" with a third round of talks involving the United States, South Korea, China, Japan and Russia.			
RS: Mr. Li mentioned this week that North Korea was ready to “step forward” on a third table of talks involving other geopolitical great powers in the region . . .			

All of the previous examples show how *MonoConc* is combined with dictionaries and a thesaurus. *MonoConc* provided quick access to collocation frequency that a thesaurus alone could not provide. Thesauri provide some examples, but not as many as *MonoConc* provides. The number of examples can show a pattern of collocation frequency useful to the users. On the other hand, if *MonoConc* does not give examples for an inquiry that a subject performs, this can also lead the users to make a decision not to use an item. At other times, subjects using *MonoConc* are looking for appropriate

sentence or phrase structures. Along with this, *MonoConc* complements English-Korean dictionary use by clarifying the part of speech of the search item. Sometimes, subjects simply want a confirmation of what they believe to be correct according to a dictionary entry. Having access to authentic examples was an additional advantage that subjects in this research study commented on because they feel that entries in dictionaries may not be reliable and could sound nonnative. Slang terms can also be found in *MonoConc*, giving nonnative speakers insight into authentic-sounding language.

This section has included the use of *MonoConc* as a complementary tool following the use of dictionaries and thesauri. The next section describes cases where *MonoConc* was used first, with other tools following in the same episode. In these examples, the *MonoConc*/English-Korean dictionary were combined because *MonoConc* did not provide enough examples. This subject also complained about *MonoConc* giving too many examples, so that she could not find what she was looking for.

In Table 4.36, the subject used *MonoConc* first and typed the search word ‘plot.’ She said that she knew it had been used as a noun in the article, but wanted to see if it could be used as a verb. She wanted to use it as a verb, ‘plotting,’ and looked at the collocation frequency. Then, she used the *Yahoo* dictionary and saw that ‘to’ comes after ‘plot’ and that it was used as a verb. She used the online dictionary because *MonoConc* had only given her 10 examples. In EK, she typed ‘plot’ and found ‘plot to kill a person.’ This example was enough information for the subject to write ‘plotting to overview.’ Although this subject saw 10 examples in *MonoConc* and said she did not see enough examples of what she was looking for, she pointed out that if she saw the right kind of example in EK, one was enough for her to make a decision based on what she thought to be correct.

Table 4.36. Subject JHC using *MonoConc* with search term "plot" and English-Korean dictionary with search term "plot"

No.	Observation	Conversant	Verbatim
1	VO Typing 'plot' in M	Researcher	'Plot.'
2		JHC	Plot' was used as a noun in the article, but I wanted to see if it could be used as a verb. I thought the <i>Yahoo</i> dictionary was better for this, so I closed the window and used the dictionary.
3		Researcher	Why did you look for it as a verb?
4		JHC	I wanted to use it as a verb. I wanted to write about plotting, but 'plot' was used as a noun in the article.
5	VO Looking at CF	Researcher	You were looking at CF.
6		JHC	I then used <i>Yahoo</i> (dictionary).
7		Researcher	Why did you look for in CF?
8		JHC	To see if it could be used as a verb.
9	AC Pointing at 'to' most frequent word 1-right	JHC	'To' comes after 'plot,' and it means that 'plot' was used as a verb.
10		Researcher	You said you were looking for 'plot' as a verb. Why did you use <i>Yahoo</i> after using M?
11		JHC	It gave me only 10 examples. I didn't see a lot of examples. I didn't see examples that I was looking for. So, I used <i>Yahoo</i> .
12	VO Typing 'plot' in EK	Researcher	What were you looking at?
13		JHC	I found 'plot to kill a person.' It caught my eye, not like M, which shows too many examples. I used 'plot to' because I saw it.
14	VO Typing 'plotting to overview' in W		
OS: Complete nuclear dismantling is a plot to overthrow the North's socialist system after stripping it of its nuclear deterrent, it said. Verifiable nuclear dismantling reflects a U.S. intention to spy on our military capabilities before starting a war," it also said.			
RS: North Korea accused America of plotting to overthrow North Korean social system, dismantling military capabilities, and eliminating peaceful nuclear energy industry.			

After using *MonoConc*, the next subject (see Table 4.37) used the *MonoConc* and KE combination of tools. He used *MonoConc* first to find out if the that-clause comes

after the verb ‘criticize.’ He only found ‘for + noun phrase’ examples following ‘criticize.’ He decided to use another word with the same meaning in English, so he used a Korean word ‘binan,’ which means criticize in English. The purpose of looking for this word in KE was to find the translation of that word in English (see lines 7 & 8). He chose ‘reproach,’ hoping that it would be followed by a that-clause and then went back to *MonoConc* in order to see what came after ‘reproach.’ Again, he found that ‘reproach’ was followed by ‘for + noun phrase.’ He decided to use neither ‘criticize’ nor ‘reproach.’ Instead, he used the verb ‘say,’ which he knew could be followed by a that-clause (see Table 4.37). This example shows that these two tools complement one another. KE complements *MonoConc* by helping find a correct search word, which was a synonym. On the other hand, *MonoConc* complements KE by providing examples of the search word.

In the following example (see Table 4.38), the subject used *MonoConc* and EK as his combination of tools. *MonoConc* was used first because he did not know the meaning of ‘fitful.’ He wanted to find the concept of ‘fit’ in *MonoConc*, and thought that ‘fitful’ could be the adjective form of ‘fit.’ His theory was that ‘fitful’ could be a combination of ‘fit + ful’ and that the meaning could be ‘appropriate.’ He wanted to see how it was used in sentences and saw that the word ‘dreams’ came after it. He knew that ‘appropriate dream’ did not make sense (see line 2). However, he found that this theory was not correct and decided to use a dictionary. In EK, he looked up the definition in the dictionary because he found in the corpus that his guess had been wrong and he needed a clear definition of the word ‘fitful.’ During this tool sequence, *MonoConc* helped the subject understand the meaning of a word by providing a context, which was further explained in EK.

Table 4.37. Subject KWA using *MonoConc* with search term "criticiz*" and Korean-English dictionary with search term "binan"

No.	Observation	Conversant	Verbatim
1	VO Typing 'criticiz*' in M	Researcher	You came back to M and typed 'criticize.' What were you thinking when you typed this?
2		KWA	North Korea may be the subject. I wanted to say that North Korea is criticizing the United States. But I wanted to use a clause because there is a lot to explain. I wasn't sure if a clause could follow 'criticize.' I thought it could, but I didn't see any examples with clauses like this.
3		Researcher	You mean using a that-clause?
4		KWA	Yes. By using a that-clause, I wanted to describe what the United States did. I was looking for examples like this. But I only found examples that have 'for' followed by a noun phrase. All examples show that a noun phrase is used after 'criticize.' Eventually, I didn't use 'criticize.' Or, did I use it? Anyway, I just wanted to find out if 'for' or 'that' was used after 'criticize.'
5	VO Typing 'binan' (criticize in English) in KE	Researcher	You didn't see any that-clauses, so what decision did you make?
6		KWA	I don't remember.
7		Researcher	Let's watch it in the video later. Why were you looking for 'criticize' in KE?
8		KWA	I couldn't find 'criticize' with a that-clause in M. So, I was looking for different words that are used with that-clauses. So I chose 'reproach' and looked for it in M later.
9	VO Typing 'reproach' in M	Researcher	Why did you choose it?
10		KWA	I was wondering if 'that' comes after 'reproach.' If it could, I was going to use it. But I don't think I'll find it in M.
11		Researcher	You are looking for 'reproach' in M.
12		KWA	Yes. In the example, 'for' is used with noun phrases.
13	VO M is displaying results.	Researcher	What else are you looking at?
14		KWA	I was waiting for more examples to show up, but I never saw examples with that-clauses.
15		Researcher	Let's see what you did.

Table 4.37. Continued

16	KWA	I didn't use 'criticize.' Instead, I used 'say' because 'say + that-clause' appeared a lot. So, I just decided to use it.
OS: "The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime," it said, using the initials of North Korea's official name, the Democratic People's Republic of Korea.		
RS: A North Korea's official said that the present policy of the U.S. just makes the Korean peninsula in danger, which can occur a nuclear war.		

Table 4.38. Subject MKS using *MonoConc* with search term "fitful" and English-Korean dictionary with search term "fitful"

No.	Observation	Conversant	Verbatim
1	VO Typing 'fitful' in M	Researcher	What were you thinking when you looked for 'fitful' in M?
2		MKS	First of all, I didn't know the meaning. I knew the meaning of 'fit' was 'appropriate.' I wanted to see how it was used in sentences. I saw 'dream' came after it, so it's not 'appropriate dream.' And I used a dictionary. The reason I used M before using a dictionary was to know the concept of 'fit.' I thought 'fitful' might be an adjective form of 'fit.' I thought 'fitful' was made from 'fit' + '-ful.' So, the meaning is 'appropriate.' I guessed that way at first. But my guess was wrong. So, I used a dictionary.
3		Researcher	Did you see any examples other than 'dream'?
4		MKS	I saw 'sleep.' I just looked at it to see whether my guess was correct or not. In this case, I didn't get much out of it. If I had known the meaning before using it, M might have helped a lot.
5	VO Typing 'fitful' in EK	Researcher	So, you were looking for 'fitful' in EK.
6		MKS	I needed to know the definition. I didn't use the dictionary not because the corpus didn't give the answer, because I found in the corpus that my guess was wrong.
OS: . . . raising doubts about whether the fitful negotiations are making even limited progress.			
RS: . . . registering doubts about whether the seeming negotiations are making even limited progress.			

4.4.3. Combination of English-English dictionary and

MonoConc

In the following example (see Table 4.39), the subject used *MonoConc* along with EE to understand the meaning of the phrase ‘put forward.’ First, she used *MonoConc* in order to find out the meaning of the phrase by reading the sample sentences. She typed ‘put @ forward’ in *MonoConc* and found out that the results were not what she was expecting to see. She changed her search phrase to ‘put forward’ and clicked on some sentences. She tried to understand the phrase by reading the whole sentence. They were similar to what she was looking for, so she used a dictionary to confirm that she was right. Then she used the search phrase ‘put forward’ in EE, but did not get any results because it is a phrase. Her next attempt was searching for ‘put’ in EE, but did not find any phrases with ‘put’ and ‘forward’ used together. After this she used the word ‘try’ in her writing because she thought that is what ‘put forward’ meant after reading the results in *MonoConc* (see line 21).

Through such a sequence of tool use, *MonoConc* complements the dictionary by permitting the user to enter phrases, which is not possible in dictionaries. Another way *MonoConc* is a complementary tool is that it gives easy access to sample sentences with flexible search functions. This could involve several steps: deciding on a search word, looking for the search word, finding the phrase you are looking for, and reading the sample sentences if they are provided. By using special search functions such as * and @, *MonoConc* allows both specific and broad search options.

Table 4.39. Subject JWK using *MonoConc* with search term "put@forward" and English-English with search term "put forward"

No.	Observation	Conversant	Verbatim
1	VO Typing 'put@forward' in M	Researcher	Why did you look for 'put @ forward' in M?
2		JWK	I was wondering if this phrase was used with different meanings in other sentences. I also used dictionaries to find the definition. Because I used @, I didn't find what I wanted. 'Put' and 'forward' should be used side by side. When there was a word in between the two words, the results were not what I wanted. So, I looked for 'put forward' without @.
3	VO M showing results of 'put @ forward'	Researcher	Did you mean you didn't like the results?
4		JWK	No. I shouldn't have put @. In the news article, it's 'put forward.' So, you deleted @ and only typed in 'put forward.' Why did you do that? I was looking to find out the meaning, to see how it was used in different sentences, and to see if the meaning I had in mind was correct.
5		Researcher	What were you reading?
6		JWK	I read each sentence as a whole by clicking on them.
7		Researcher	Did you find what you were looking for?
8	VO Clicking sentences and reading them in M	JWK	I thought they were similar to what I thought. I used a dictionary to make sure.
9		JWK	Here, I knew the meaning of 'put forward.' I just wanted to make sure by using a dictionary, so I know the meaning in Korean and in English.
10		JWK	But I couldn't find it in the dictionary.
11	VO Typing 'put forward' in EE	JWK	Because it was a phrase, so it didn't give me any results.
12	VO Showing no results	JWK	I didn't find anything, so I just used what I saw.
13		Researcher	What you saw?
14		JWK	I used M and got the meaning, so I paraphrased the phrase according to the meaning.
15	VO Typing 'put' in EE	JWK	I looked for 'put' here but didn't find anything.
16		Researcher	What were you looking for in EE?

Table 4.39. Continued

17		JWK	I was looking for a phrase but didn't know how to search for a phrase.
18	VO Typing 'the US has tried to finish' in W	Researcher	What did you use instead of 'put forward'?
19		JWK	Try.
20		Researcher	How did you come up with that word?
21		JWK	I thought that 'put forward' meant to try, so I used 'try.' I was going to find synonyms of 'put forward' because I didn't want to change the meaning. Since I wasn't able to find any, I translated the meaning I had in mind in Korean into English.

OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs, raising doubts about whether the fitful negotiations are making even limited progress.

RS: The clear rejection by North Korean radio on Saturday the formula which the U.S. has basically tried to finish negotiations about the North Korea's nuclear program is increasing skepticism about whether the intermittent talks has advanced in a small portion.

4.4.4. Combination of *MonoConc* and *MonoConc*

In order to get more accurate and specific results from *MonoConc*, the subjects oftentimes decided to use *MonoConc* two or more times in a row within one episode with different options and variations of search words. The following excerpts are examples of subjects using *MonoConc* more than once to carry out their searches.

In the first example (Table 4.40), the subject used *MonoConc* twice in a row in order to perform specific searches. She looked for 'aim' in *MonoConc* the first time to find the collocates that follow aim. She wanted to find out how often and in what context 'aim to' and 'aim of' were used in the corpus. She examined the sample sentences and collocation frequency. After that, she thought 'aim of' did not sound appropriate in relation to her sentence (see line 4). She then looked for 'aim to' in order to see more examples that included those two words specifically. When she only typed 'aim' in

MonoConc, she was able to find only one example of ‘aim to.’ By using a more specific search term, she could narrow the search and get the results that she wanted. After examining the results, she found similar patterns that she could apply to her sentence. This example demonstrates a series of entries in *MonoConc* in which she moves from general searches to more specific searches.

In the following example (see Table 4.41), the subject performed two searches with *MonoConc*. She compared her two searches and made a decision. First, she looked for ‘announce* that’ in *MonoConc*. She wanted to see if she could use ‘announce to.’ She was going to write ‘announce that’ at first, but saw that ‘announce to’ was used more often. She compared the frequency of ‘announce that’ and ‘announce to.’ Next, she typed ‘announce* to’ in *MonoConc* (see line 5) and saw that a noun often came after it. She paid careful attention to the frequency of nouns following ‘announce to’ and if the context of the sentence was similar to what she was going to write. When she compared ‘announce that’ and ‘announce to,’ she saw that ‘announce to’ was usually followed by a person or a noun, but not a verb. This information enabled her to make a decision and use a that-clause. In order to confirm her answer, she checked the collocation frequency of ‘announce* to’ to see if a verb or noun followed ‘to’ more frequently. Because there were not many examples in which ‘announce that’ was followed by a verb, she used a that-clause. She also thought that it sounded better and she could simply write what she was going to write. In this episode, the two separate *MonoConc* searches led the subject to find which expression to choose based on the frequency and part of speech of the words she was planning to use with the search word. In this case, *MonoConc* gave her the opportunity to compare two different verb phrase structures. By analyzing those two structures, the subject was able to refine her writing.

Table 4.40. Subject KYL using *MonoConc* with search term "aim"

No.	Observation	Conversant	Verbatim
1	VO Typing	Researcher	Why were you looking for 'aim' in a corpus?
2	'aim' in M AC Pointing at the article	KYL	Here, it says 'aimed at.' I thought it's usually 'aim to.' I wanted to use 'aim to.' I wanted to see how it's used.
3	VO Typing	Researcher	How did you decide to use 'of'?
4	'aim of' in W	KYL	I didn't use 'to' because of a gerund (verb-ing) coming after it. At first, I used 'of,' but it sounded awkward. Then, I changed it to 'aim to end.' I wanted to make sure that 'aim to' could be used, but there was a noun after it. So, I changed it to 'aim of,' but I didn't like it. It doesn't seem correct to write 'of ending' because there is 'is aimed' in front of it. So, I changed it to 'aim to.' I knew 'aim to' and wanted to be sure.
5	VO Looking at CF	Researcher	You were looking at the collocation frequency.
6	AC Pointing at number 11, the frequency of 'of' as the most frequent collocation	KYL KYL	I used 'of' because of 'ending.' And I changed it to 'to.' I thought 'of ending' sounded awkward. I knew 'aim to' and wanted to see examples of it. When I looked for it, I saw 'aim of,' 'aim at,' and 'aim to.' I used 'of' because 'ending' came after 'of.' But it was awkward, so I wanted to change it to 'aim to.' That's why I looked for it in CF to see how often 'aim of' and 'aim to' were used, and to find out if I could use 'aim to.' Here, it's used 11 times.
7	VO Typing 'aim to' in M	KYL	'Of' is used a lot, so I thought it's ok to use 'of.' But I didn't want to use 'of,' so I wanted to see how 'aim to' was used in a sentence. I also wanted to find out if I could use it for my sentence. Last time, I found one example with 'aim to.' I wanted to see more examples. Actually, there are a lot of examples.
8		Researcher	Here are the results of 'aim to.' What were you looking at?
9	VO Changing 'aimed of ending' to 'aimed to end' in W	KYL	A verb came after 'aim to.' First, I saw how often 'aim to' was used. Then, I found out that a verb came after 'to,' so I changed it to 'to end.'
10		KYL	A noun could come after a preposition, but there is a verb in this case. I changed 'ending' to 'end.' First, I wrote 'ending of,' but I didn't like it. So, I looked for 'aim to' to see how often it was used, and to see if it's correct. I found 'aim to + a verb' in the results, so I changed it to 'aim to.'
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: North Korean radio announced on Saturday that they rejected the formula the United State suggested as its pragmatic position in conversation which is aimed to end North Korea's nuclear programs.			

Table 4.41. Subject JYP using *MonoConc* with search term "announc*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'announc* that' in M	Researcher	What were you thinking when you looked for 'announce' in M?
2		JYP	I wanted to know if I could use 'announce to.' At first, I was going to write 'announce that.' I wanted to see if I could use 'announce to.' It seemed that 'announce that' was used more.
3		Researcher	What were you looking at?
4		JYP	I was looking at how many times 'announce that' was used and compared the frequency with that of 'to.'
5	VO Typing 'announc* to' in M	JYP	This time, you were looking for 'announce to.' There were more cases that a noun came after it. I was wondering whether to use 'announce to' or an 'announce that-clause.'
6		Researcher	What did you pay attention to?
7		JYP	If a noun followed more frequently, and if the context of the examples were similar to the context of the sentence I was going to write it.
8		Researcher	What were you comparing between 'announce that' and 'announce to'?
9		JYP	'Announce to' is usually followed by a person or a noun, but not by a verb. So, I decided to use a that-clause.
10		Researcher	You already wrote 'announce that' in W.
11		JYP	I was just checking.
12		Researcher	Why were you looking at the frequency?
13	VO Looking at CF of 'announc* to' in M	JYP	To see if a verb or noun came after 'to' more frequently. There weren't many cases that 'announce that' was followed by a verb, so I used a that-clause. 'Announce + that-clause' was used more frequently, and it sounded better. Also, I didn't have to say to whom to announce. Instead, I was going to write what to announce.
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks . . .			
RS: North Korean radio explicitly announced that they decided not to accept the formula the United States has suggested as its basic position in talks . . .			

In the following episode (as shown in Table 4.42), the subject used *MonoConc* twice as she had done in the previous episode. She typed ‘skeptical about’ in *MonoConc* to see if it was correct and to see if she would see ‘skeptical on’ (see lines 1 and 2). She wanted to see the frequency of ‘skeptical on’ and to compare the context to that of the one she was writing. Next, she searched for ‘skeptical on’ in *MonoConc* and did not see any results, and so she cancelled the search. In this case, the lack of information in *MonoConc* led the user to assume that the phrase she was going to use was not correct: she was therefore still able to make a decision by using *MonoConc*.

Table 4.42. Subject JYP using *MonoConc* with search term "skeptical"

No.	Observation	Conversant	Verbatim
1	VO Typing 'skeptical about' in M	Researcher	Why did you look for 'skeptical about' in M?
2		JYP	I wrote 'skeptical about,' and was making sure if it's correct. I was also looking to see if 'skeptical on' showed up in M. I thought I was looking for 'on' also.
3		Researcher	What were you looking at?
4		JYP	I wanted to see how often it's used in sentences and if the context was similar to that of the sentence I was writing.
5	VO Typing 'skeptical on' in M	Researcher	This time, you were looking for 'skeptical on' in M.
6	VO M is searching for examples. VO Canceling the search.	JYP	If it didn't show any results within this much time, there wouldn't be any results. So, I used 'about.' I looked for both in M and found 'about' more.
OS: . . . raising doubts about whether the fitful negotiations are making even limited progress.			
RS: Because of this announcement, people are skeptical about the effectiveness of this negotiation.			

Table 4.43 shows a sequence of *MonoConc* use in which the subject obtains very specific examples by narrowing the searches with specific search phrases. The subject

looked for ‘destruction’ in *MonoConc* to change the phrase ‘stifle the country.’ While she was checking the examples, she saw the phrase ‘destruction of the country.’ She decided to look for the phrase ‘destruction @ country’ in *MonoConc* to see if it was authentic (see line 5). She saw only one example, but decided to use it anyway because she thought it would convey the correct meaning. In the previous episode, the criterion for the subject to choose to use the expression found in *MonoConc* was its high frequency. However, in this episode, the subject thought one example was enough to prove that her idea was correct.

In summary, the results show that *MonoConc* complements dictionary use and thesauri in various ways. When the subjects know the meaning and syntactical structure of the words they want to use in a sentence (as found in thesauri or Korean-English dictionaries), they do not have to use *MonoConc*. When subjects find the meaning through the use of dictionaries, but do not know how to use it in a sentence, they then combine their search with *MonoConc*. *MonoConc* is also used with dictionaries for authentic examples of verbatim language, while dictionaries provide insight to the possible meanings. When *MonoConc* is used multiple times consecutively, it is because subjects start with broad searches and narrow them down, and/or they want to compare verb phrases and the words that follow them.

Table 4.43. Subject HJS using *MonoConc* with search term "destruction"

No.	Observation	Conversant	Verbatim
1	VO Typing 'destruction' in M	Researcher	Why did you look for 'destruction' in M?
2		HJS	There is 'stifle the country' in the article. I was going to change it. I thought that 'stifle' was 'destruction of the country.' I was checking to see if it was used in real life. But I only saw different examples, so I just used 'destruction of the country.' I found 'destruction of housing.'
3		Researcher	What did you decide?
4		HJS	I looked for 'destruction of the country.'
5	VO Typing 'destruction @ country' in M		I was wondering if I could use it like this. I made it up, so I was checking to see if it was used in real life.
6	VO Typing 'the destruction of the country' in W		
7	VO M showing one example	HJS	There weren't many examples. What did you decide? It wasn't used very frequently, but I thought it was possible to use it to convey the meaning.
8	VO Typing 'destruction of the country' in W		
OS: "Irreversible nuclear dismantling is nothing other than a noose to stifle us after eradicating our peaceful nuclear-energy industry," it said.			
RS: North Korea expressed its aversion to this demand, saying that US aims to spy on its military facilities, agitate a war against North Korea and furthermore, put them trapped in the destruction of the country.			

4.5. What are the effects of using a corpus on the accuracy of learners' writing?

Research data for this question were obtained from two third-party English-native raters' judgments on the subjects' writings. The judgment was made whether an expression is native-like or not. The two raters were American English native speakers who have not been involved in any procedures of this research. These raters graded the

writing samples separately and reconciled any differences in their judgments if needed. Based on these grades, the investigation included the ratio of successful and unsuccessful episodes, sequence of consultation, results found in a corpus, categories of errors, and information (or sequence) that causes nonnative- or native-like expressions. These results then led to conclusions for training learners to solve problems caused by corpus use.

Table 4.44 shows individual subjects' use of each tool and whether the writing was correct or incorrect. The results show that when the subjects used any of the tools, they correctly applied the findings to their writing 81% of the time (101 out of 171 episodes). Eight of 10 subjects correctly applied the findings to their writing in over 80% of the total episodes. Among all of the episodes, 59% of the episodes showed that the subjects used correct wording after referring to reference tools (Figure 4.5). Fourteen percent of the episodes resulted in incorrect production of writing. Twenty-seven percent of the episodes resulted in subjects using the tools but not relating the information to their writing. This category of usage is labeled "Null" in the charts below and applies to when English-Korean and English-English dictionaries were used to look up definitions of words in the newspaper article. Sometimes the subjects looked for information and decided not to use the findings in their writing for reasons such as inappropriate context, meanings, or structure. When comparing correct and incorrect episodes, they produced correct words or phrases 81% of the time (Figure 4.6).

MonoConc was the most used tool out of all of the reference tools (Table 4.44). One hundred fifty-two times out of 321 times MonoConc was used. Out of all the tools available, MonoConc was used 47% of the time. The ratio of MonoConc use out of all tool use ranges from 25% to 83%, while a range of 27% to 100% of the time the subjects produced correct outcomes (Figure 4.7). An average of 80% of the time the outcomes were correct.

Table 4.44. Reference Tool Use and Grammaticality

Subject	Number of Episode	Tool Use (Times)					Grammaticality (Episode)			
		M	EK	KE	T	EE	Correct	Incorrect	Null	Ratio (%)
HJC	16	13	3	4	8	2	10	1	5	90.9
JDO	20	14	4	4	3	1	14	2	4	87.5
JHC	11	12	5	0	1	0	7	1	3	87.5
JWK	15	15	0	0	13	3	3	8	4	27.3
JYP	14	30	4	0	1	1	11	1	2	91.7
KWA	31	23	9	11	8	0	20	5	6	80.0
KYL	15	12	8	4	4	3	7	1	7	87.5
MAP	13	10	19	3	8	0	4	3	6	57.1
MKS	22	15	9	10	10	0	17	2	3	89.5
YJK	14	8	1	0	0	5	8	0	6	100.0
Total	171	152	62	36	56	15	101	24	46	80.8

Figure 4.5. Ratio of Grammaticality (Correct: Incorrect: Null)

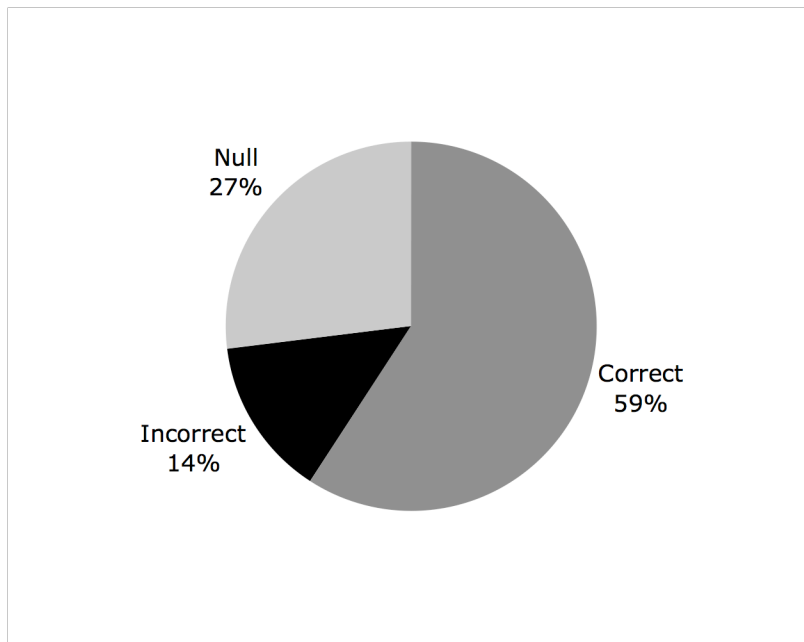
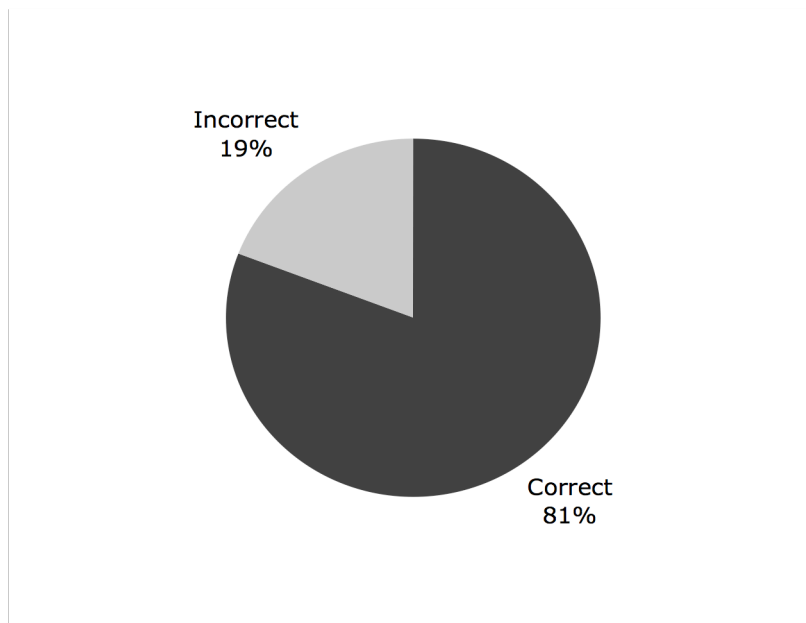
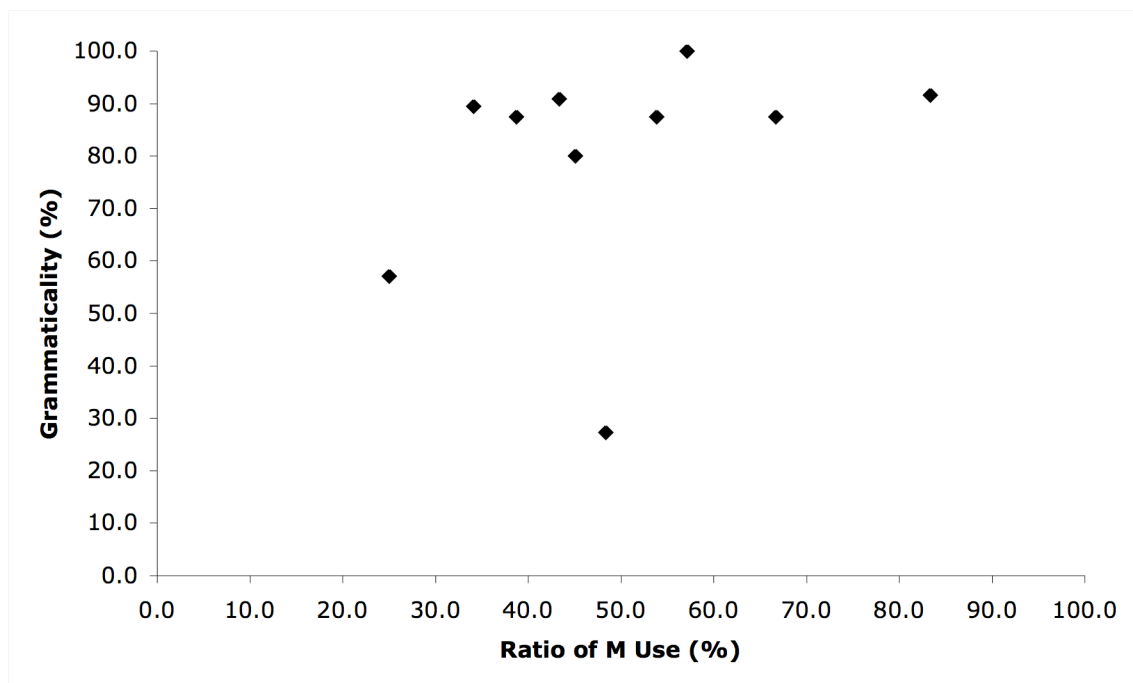


Figure 4.6. Ratio of Grammaticality (Correct: Incorrect)

Figure 4.7. Ratio of *MonoConc* Use and Grammaticality

The results show that *MonoConc* is most useful for sentence-level composition. The subjects looked for examples that had similar patterns, structure, or the exact formulation they were planning to use in their sentence. During the process of searching for similar sentences and copying the structure into their sentences, *MonoConc* can lead users to produce ungrammatical phrases or sentences. Among various advantages of using *MonoConc*, the subjects took advantage of keying in phrases as search terms, which is not a function that dictionaries provide. In this section, the advantages of using *MonoConc* and the effects it has on subjects' writing accuracy will be shown.

4.5.1. Repeated search word

4.5.1.1. Repeated search word: "blame"

This section deals with the subjects' interaction with *MonoConc* and the effects it has on their writing. The specific focus is on the similarities and differences of results in writing when using the same search word. At times, students were able to find the correct example and apply it correctly in their writing, but they were not always able to apply the patterns as seen in the examples, although their writing required a similar structure.

The subject in the following excerpt successfully used *MonoConc* as she looked for the correct expression to accompany the word 'blame.' She wanted to change the word 'accusing' to 'blame,' but was not sure if 'blame + person + for' or 'blame for' was correct. She found what she was looking for and wrote 'blame the United States for using.' She used the phrase correctly in her writing and was able to confirm that her guess was correct. (See Table 4.45 for details on this episode.)

Table 4.45. Subject JDO using *MonoConc* with search term "blame"

No.	Observation	Conversant	Verbatim
1	VO Typing 'blame' in M	Researcher	What were you thinking when you looked for 'blame' in M?
2		JDO	I changed 'accusing' to 'blame.'
3	AC Reading the article	JDO	Here is 'accuse A of B.' I knew 'blame' was followed by 'for.' But, I wasn't sure whether it's 'blame for' or 'blame + person + for.' I was looking for the structure.
4	VO Closing the result window quickly	Researcher	You cancelled the search and closed the window immediately. Did you find what you were looking for?
5	AC Pointing at the results	JDO	Yes, I did. Here is 'for.' I was checking if I could write 'blame the United States for using' or 'blame for the United States using.' I knew it before and found examples.
6	VO Typing 'blame United States for planning a war' in W		
OS: It used typically unrestrained language in accusing the United States of secretly planning a war.			
RS: The radio broadcast blamed United States for planning a war, as usually, using offensive languages.			

The next subject, JYP (see Table 4.46), used *MonoConc* to look up a phrase that she needed to know. She performed three searches with variations of phrases using the word 'blame.' She did not know whether 'blame' was followed by 'for' or if it was used with a that-clause. She also wanted to see the frequency. She also entered 'blam* for.' This time she utilized the @ character and saw numerous examples. Her third entry was 'blam*@for' and found more examples (see line 10). Finally, she was able to make a decision and typed 'blamed the US Hawks and followers for the danger.' *MonoConc* enabled her to make inquiries on what she thought could possibly be correct.

Table 4.46. Subject JYP using *MonoConc* with search term "blame* that"

No.	Observation	Conversant	Verbatim
1	VO Typing 'blame* that' in M	Researcher	Why were you looking for 'blame* that' in M?
2		JYP	Whether 'blame' could be followed by 'for' like 'accuse,' and if 'blame + that-clause' was used. I was also looking for which expression was used more frequently.
3		Researcher	What did you find?
4	VO Showing 6 results in M. AC Pointing at 'He blames that same negativity . . . ' in M	JYP	This 'that' is not a part of a that-clause. It's a part of 'that same.' So, I thought 'blame + that-clause' was not used. I looked for 'blame + a person + for.'
5		Researcher	Did you not get what you want?
6		JYP	I was just checking. I was looking to see how to use 'blame that' and 'blame for.' But 'blame that' was not used.
7	VO Typing 'blam* for' in M	JYP	There were a lot of examples with 'blame for,' so I thought about using it.
8		Researcher	What were you looking at?
9		JYP	Reasons for blaming come after the verb 'blame.'
10	VO Typing 'blam*@for' in M	JYP	I was also looking for 'blame + a person + for,' and I found some examples. Previously, I had looked for 'blame for.' I was wondering if I could put a person in between. According to the examples, it's possible.
11	VO Typing 'blamed the US Hawks and followers for the danger' in W	JYP	So, I thought I could write 'blamed the US Hawks and followers for.'
OS: The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K.			
RS: North Korea blamed the U.S. hawks and followers for the danger of the Korean Peninsular.			

Another subject, KWA, also searched for 'blame' in *MonoConc*. He only typed 'blame' and checked the collocation frequency of 'for.' He saw that it was the second-most-frequent word. He did not see any examples using that-clauses, so he decided to

write 'blame for' in his paraphrase. This subject may have put himself at a disadvantage by only entering the word 'blame' in *MonoConc* and not using search characters or by typing in variations of what he thought may be correct. (See Table 4.47 for the outcome of this episode.)

Table 4.47. Subject KWA using *MonoConc* with search term "blame"

No.	Observation	Conversant	Verbatim
1	VO Typing 'blame' in M	Researcher	Why are you looking for 'blame' in M?
2		KWA	I know 'blame for,' but I wasn't sure if it was correct.
3		Researcher	So, what did you focus on?
4		KWA	I wanted to know if there are a lot of examples that have 'for' after 'blame.'
5	VO Looking at CF	KWA	The reason I am looking at the frequency is to see if 'for' comes after 'blame.'
6	VO 'For' appears as the second most frequent word to the right.	KWA	I'm looking at 'for' as the second most frequent word.
7		Researcher	Why are you focusing on 'blame for'?
8		KWA	I was thinking about whether I could use a clause. If you look at this sentence, only one word comes after 'blame for.' I wanted to know if I could use a noun clause like 'blame that . . .' I was struggling with whether 'that' comes after 'blame' or not. Originally, I wanted to put 'that,' but I couldn't find any sentences written that way.
9	VO Typing 'blame for' in W	Researcher	Why did you decide to use 'for'?
10		KWA	I didn't find any sentences with a that-clause, so I wrote 'for' after 'blame,' followed by a noun phrase. I'm trying to use a clause with a phrase.
OS: North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs . . .			
RS: The government of North Korea blamed for the policy the United States insists on giving up their nuclear programs through their radio broadcasting.			

Table 4.48 shows a second episode that the previous subject, KWA, performed in *MonoConc*. This time he began his search with the word ‘blame’ again, but started the search in the thesaurus (see Table 4.48). Because he had already used ‘blame,’ he wanted to find a synonym. He thought of ‘said,’ and then looked for the word ‘condemn’ in *MonoConc*. He still wanted to use a that-clause, so in order to get more results, he typed ‘condemn*’ in *MonoConc* in order to see more examples with verb-ing and in the past-tense (see line 5). Finally, he found a that-clause and finally decided to use ‘condemn that’ in his paraphrase. *MonoConc* helped him fine-tune his search in order to choose the most appropriate expression, albeit an incorrect one.

JDO and JYP had in mind that the verb ‘blame’ could be followed by ‘person + for.’ They both looked for examples of this phrase structure. Since they were aware of the need to write ‘person’ before ‘for,’ they were able to apply the structure they saw in *MonoConc* into their paraphrases: they were both able to use ‘blame + person + for’ correctly. On the other hand, KWA focused on finding out if he could use a that-clause with ‘blame’ and what comes after ‘for.’ When he found out ‘for’ was the most frequently occurring second word to the right, he did not seem to realize that this verb needed to have an object. As a result, he used the structure ‘blame + for’ without an object in between them in his writing. On the other hand, he found out that a noun came after ‘blame for,’ so he was able to correctly put the reason for the blame after ‘for.’ Later, he considered using the word ‘condemn’ when he was not successful in finding synonyms in the thesaurus. Unlike ‘blame,’ he found examples that showed a noun clause that directly followed ‘condemn,’ in addition to other variations. However, he incorrectly applied a different structure in his writing: he wrote “. . . North Korea condemned that the Bush administration will not achieve their main goals if they persist on their argument like the first two talks.” His sentence structure would have been correct had he not tried to make a that-clause follow the word ‘condemn.’

Table 4.48. Subject KWA using thesaurus with search term “blame” and *MonoConc* with search term "condemn"

No.	Observation	Conversant	Verbatim
1	VO Typing 'blame' in T	Researcher	You are looking for 'blame' in T.
2		KWA	I wanted to use a that-clause after 'blame.' I thought I had to use 'blame' again but didn't find any words that I liked to replace it. So, I used 'said.' I wanted to use a that-clause after 'blame.' I thought I had to use 'blame' again but didn't find any words that I liked to replace it. So, I used 'said.'
3	VO Typing 'condemn' in M	Researcher	You are looking for 'condemn' in M.
4		KWA	This word just popped into my mind because I was thinking about 'blame.' I wanted to see if 'that' could come after 'condemn.' I didn't want to use 'criticize' instead of 'blame.' 'Blame' was followed by a noun or 'for' and a noun. 'Criticize' is used in the same way. I was wondering if 'condemn' is followed by a that-clause. I was looking for a sentence with 'condemn' followed by 'that.'
5	VO Scrolling down results in M, VO Typing 'condemn*' in M	KWA	It gave me few sample sentences, so I used '*' in the search word. Then it would give me more examples with a verb-ing form and past-tense.
6	VO Scrolling down results in M	KWA	I am scanning for examples with a that-clause, but I only found those followed by nouns.
7	VO Scrolling down results in M	KWA	I was scanning for a that-clause.
8	VO Reading 'to condemn the overmedicating of the' in M	KWA	I was reading this because I thought there was a clause after 'condemn.' But it was just a noun.
9	VO Reading 'many birds condemned they were filling up'	KWA	I found this. 'Condemn' is followed by 'they were filling.' So, I found out that it was possible for 'condemn' to be used with a that-clause.
10	VO Reading 'We condemn that defamatory and' in M	KWA	I saw it and thought I could use a that-clause.
11	VO Typing 'condemn that' in W	KWA	I decided to use it.

Table 4.48. Continued

OS: While North Korea often harshly criticizes the United States for what it considers an inflexible stance, the Saturday announcement seemed to go further. It put North Korea on record as saying that it could not accept the main goals President Bush and his negotiators have insisted on in the first two rounds of talks.

RS: In the meanwhile, North Korea condemned that the Bush administration will not achieve their main goals if they persist on their argument like the first two talks.

4.5.1.2. Repeated search word: “criticize”

The second search word that several subjects looked for was ‘criticize.’ The subject JHC in this episode typed ‘criticize*’ in *MonoConc* in order to find the correct phrase. She saw the word ‘for’ in many examples. She saw the phrase ‘criticized the attorney general for’ and noticed that it was similar to her sentence and decided to use it. However, she closed her window too quickly and had to look for ‘criticize*’ again. An advantage of *MonoConc* is that if one needs to return to the examples, it is possible to re-enter the same phrase and review all examples. Her final answer was ‘criticized America for having inflexible positions,’ which would have been closer to the original had she said ‘criticized America for having an inflexible position.’ The structure of the phrase, however, was correct (episode shown in detail in Table 4.49).

Another subject (refer to Table 4.50) was searching for the correct formation for ‘criticism.’ First, she looked for ‘made criticism’ in *MonoConc* and did not get any results. Next, she tried ‘did a criticism’ in *MonoConc* and still got no results. She then tried ‘made criticisms’ in *MonoConc* and did not find any examples, so she concluded that ‘criticism’ was not used in that way. She wanted to use the phrase ‘Criticism North Korea has made’ and still thought that ‘make criticism’ was correct. Her fourth search in this episode was ‘criticism about’ in *MonoConc*, so that she could verify if ‘about’ or ‘on’ followed ‘criticism’ (see line 5). The following search phrase used in *MonoConc* was ‘criticism on,’ followed by ‘criticism over.’ At this point, she still believed that ‘about’

was a more appropriate choice than ‘over,’ but she had not made a final decision yet. Later in her writing, she arrived at: “. . . one of those typical criticisms of North Korea over an ‘inflexible stance’ of the US.” A native speaker may have instead chose the phrase ‘. . . criticisms by North Korea . . .’ but ‘of’ is also an acceptable choice. Her final product sounded sophisticated because she performed a series of searches on possible expressions and was able to find enough evidence to make a correct judgment and use ‘over’ as the preposition that should follow the object ‘North Korea.’

Table 4.49. Subject JHC using *MonoConc* with search term "criticize*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'criticized America for' in W. Typing 'criticize*' in M	Researcher	You were looking for 'criticize' in M. Why did you look for it?
2		JHC	I was wondering whether 'criticize + for' or 'criticize + America + for' was correct. I thought 'criticize + for' was correct. I was also wondering if 'criticize + of' was correct.
3		Researcher	What were you looking at?
4	AC Pointing at 'criticized the attorney general for' in M	JHC	I skimmed. I saw 'for' a lot. Here, 'criticized the Attorney General for' was used. The structure is similar to that of my sentence. It's similar to 'America for,' so I wrote this.
5	VO Typing 'criticize*' in M again	Researcher	Why did you look for it again?
6		JHC	I closed the window too early, so I wanted to see it again.
7	VO Deleting 'for' in W	Researcher	You deleted 'for' and looked at M again.
8	VO Typing 'criticized America for having inflexible positions' in W	JHC	I thought about using a noun like 'criticize American policy.' But I used 'for' again. I was trying to write it correctly.
OS: While North Korea often harshly criticizes the United States for what it considers an inflexible stance, the Saturday announcement seemed to go further.			
RS: Although North Korea had criticized America for having inflexible positions toward nuclear program, this statement went further.			

Table 4.50. Subject JYP using *MonoConc* with search term "criticism"

No.	Observation	Conversant	Verbatim
1	VO Typing 'made criticism' in M	Researcher	You were looking for 'made criticism' in M.
2	VO No results showing VO Typing 'did a criticism' in M	JYP	I wasn't sure if I could write 'make criticism.' There weren't any results. So, I was wondering if 'do' could be used like 'did criticism.' No results showed, so I gave up.
3	VO Typing 'made criticisms' in M	JYP	This time, I looked for 'made criticisms' but couldn't find any results. So, I thought 'criticism' was not used that way. I wanted to write 'Criticism North Korea has made.' Then 'make criticism' has to be a correct expression.
4		JYP	I think I wrote 'criticisms of North Korea about' because I wasn't sure if 'make criticism' was correct. I was not sure about 'did criticism' either. So, I just used a noun phrase.
5	VO Typing 'criticism about' in M	JYP	I wanted to know if 'about' came after 'criticism,' or 'on' came after 'criticism.' I was wondering which one sounds better. North Korea was the one who criticized the US's policy.
6	VO Typing 'criticism on' in M	JYP	I was wondering if I could write 'criticism of North Korea on US's stance.'
7	VO Typing 'criticism over' in M	Researcher	So, you were looking for 'criticism over.'
8		JYP	I was wondering if 'over' could be used with it. Sometimes, I don't know the differences when using different prepositions.
9	VO Looking at CF of 'criticism over' in M	Researcher	Why were you looking at the frequency of 'criticism over'?
10	VO Reading words in 2-right	JYP	To see what kind of words come after 'criticism over.'
11		JYP	There were a lot of words about government. I thought 'about' was more appropriate than 'over.'
12		Researcher	How did you decide that?
13	AC Reading	JYP	Because they criticized 'about' something, not criticized 'over' something.
14s	'Criticize Korea over the US inflexible stance' in W		I didn't change it to 'about' yet.

OS: While North Korea often harshly criticizes the United States for what it considers an inflexible stance, the Saturday announcement seemed to go further.

RS: The Saturday announcement cannot be considered one of those typical criticisms of North Korea over an “inflexible stance” of the US.

Table 4.51. Subject KWA using *MonoConc* with search term "criticiz*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'criticiz*' in M	Researcher	You came back to M and typed 'criticize.' What were you thinking when you typed this?
2		KWA	North Korea may be the subject. I wanted to say that North Korea is criticizing the United States. But I wanted to use a clause because there is a lot to explain. I wasn't sure if a clause could follow 'criticize.' I thought it could, but I didn't see any examples with clauses like this.
3		Researcher	You mean using a that-clause?
4		KWA	Yes. By using a that-clause, I wanted to describe what the United States did. I was looking for examples like this. But I only found examples that have 'for' followed by a noun phrase. All examples show that a noun phrase is used after 'criticize.' Eventually, I didn't use 'criticize.' Or, did I use it? Anyway, I just wanted to find out if 'for' or 'that' was used after 'criticize.'
5	VO Typing 'binan' (criticize in English) in KE	Researcher	You didn't see any that-clauses, so what decision did you make?
6		KWA	I don't remember.
7		Researcher	Let's watch it in the video later. Why were you looking for 'criticize' in KE?
8		KWA	I couldn't find 'criticize' with a that-clause in M. So, I was looking for different words that are used with that-clauses. So I chose 'reproach' and looked for it in M later.
9		Researcher	Why did you choose it?
10		KWA	I was wondering if 'that' comes after 'reproach.' If it could, I was going to use it. But I don't think I'll find it in M.
11	VO Typing 'reproach' in M	Researcher	You are looking for 'reproach' in M.
12		KWA	Yes. In the example, 'for' is used with noun phrases.
13	VO M is displaying results	Researcher	What else are you looking at?
14		KWA	I was waiting for more examples to show up, but I never saw examples with that-clauses.
15		Researcher	Let's see what you did.

Table 4. 51. Continued

16	KWA	I didn't use 'criticize.' Instead, I used 'say' because 'say + that-clause' appeared a lot. So, I just decided to use it.
OS: "The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime," it said, using the initials of North Korea's official name, the Democratic People's Republic of Korea.		
RS: A North Korea's official said that the present policy of the U.S. just makes the Korean peninsula in danger, which can occur a nuclear war.		

In the case of the participant shown in Table 4.51 shown above, *MonoConc* provided examples that helped him determine which structure he should use for his sentence. KWA actually found the same structure in the samples and used the same structure in his writing. In a similar way, JYP (see Table 4.50) was also looking for collocates after the noun form of the verb 'criticize.' In her case, *MonoConc* provided examples that allowed her to compare the use of different prepositions. KWA (Table 4.51) was persistent in wanting to use a that-clause after the verb 'criticize' but found examples that showed 'criticize for + noun phrase.' In this case *MonoConc* displayed examples that proved his theory was incorrect, so he decided to use another verb that was suitable to his theory. As a result, *MonoConc* prevented him from using a wrong verb phrase structure in this episode.

4.5.1.3. Repeated search word: "accuse"

This section concentrates on how four subjects used the search word 'accuse' in *MonoConc*, the conclusions they drew from its examples, and how they applied what they learned in their paraphrases. The first subject typed 'accused' in *MonoConc* (see Table 4.52). First, she thought that she could use 'excoriate,' but finally gave up and decided to use 'accused.' She wanted to see how it was used in sentences and found out that the structure she could use was 'accuse + person + of,' so she wrote 'accused the

United States of . . . ' Her final paraphrase stated: “ . . . North Korea accused the US of taking a strict stance . . . ” which shows that she chose the correct structure for this particular expression.

Table 4.52. Subject JDO using *MonoConc* with search term "accused"

No.	Observation	Conversant	Verbatim
1	VO Typing 'accused' in M	Researcher	It's a different search. You were looking for 'accused' in M.
2		JDO	I thought that I could use 'accused.' I gave up using 'excoriate.'
3		Researcher	What were you looking at?
4		JDO	Based on the previous search, I learned that 'excoriate' has the meaning of 'criticize.' I wanted to see how 'excoriate' was used in a sentence, but there weren't any examples. So, I wanted to use 'accused' because I knew it and felt comfortable using this word. I was checking to see how it was used because this word was usually used like 'accused of.' So, I wanted to see how it was used in sentences.
5	VO Typing 'accused' in front of 'excoriate' in W	Researcher	What did you find out?
6		JDO	'Accuse + person + of.' 'Accused Giuliani of.' I found 'someone of something.' So, I wrote 'accused the United States of blah blah blah.'
7		Researcher	Let's see.
8		JDO	First, I used 'excoriate,' but there weren't any examples. Here, I was looking for what was being accused. In order to use 'accuse A of B,' I changed 'inflexible' to 'strict.' I used 'strict stance' to replace 'inflexible stance.'
OS: While North Korea often harshly criticizes the United States for what it considers an inflexible stance, the Saturday announcement seemed to go further.			
RS: On occasion, North Korea accused the US of taking a strict stance, and the Saturday announcement allures that North Korea will go further.			

In Table 4.53, the participant typed ‘accuse*’ in *MonoConc* because she saw the phrase ‘accused of’ in the article, and wanted to know whether ‘accused of America’ or ‘accused America of’ was correct. She saw the structure ‘accused + something + of’ in *MonoConc* and looked for patterns in collocation frequency. The video observation of her activity in *MonoConc* showed that ‘of’ was the most frequent word as the first to the right and second to the right. She saw that ‘of’ occurred even second-most frequently (see line 6) as the 2-right word, so she correctly typed: “North Korea accused America of plotting to overthrow . . .” in her paraphrase.

Table 4.53. Subject JHC using *MonoConc* with search term "accuse*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'North Korea accused America' in W. Typing 'accuse*' in M	Researcher	Why did you look for 'accuse*' in M?
2		JHC	I saw this in the article. 'Accused of' was in the article. I wanted to know whether 'accused of America' or 'accused America of' was correct. That's what I was looking for.
3		Researcher	What were you reading?
4		JHC	'Accused + something + of.' I didn't read a specific sentence. I looked for patterns.
5	VO Looking at CF	Researcher	You were looking at CF.
6	VO M showing 'of' as most frequent word as 1-right and 2-right)	JHC	'Of' occurred a lot. There were no other prepositions other than 'of.' 'Of' occurred even second most frequently as the 2-right word, so I thought I could put America in between.
7	VO Typing 'America of' in W		
OS: Complete nuclear dismantling is a plot to overthrow the North's socialist system after stripping it of its nuclear deterrent, it said.			
RS: North Korea accused America of plotting to overthrow North Korean social system, dismantling military capabilities, and eliminating peaceful nuclear energy industry.			

The third subject began by using ‘accus* @ of’ as her search phrase in *MonoConc*. She wanted to find expressions using the structure ‘accuse + person + of’ and saw an example. She also used the *Yahoo* dictionary to look for examples. Next, she typed ‘accus* that’ in *MonoConc* and saw that most examples used ‘accusation.’ She then decided to use the structure ‘accuse + a person + verb-ing.’ The researcher pointed out that previously, she had gotten results with ‘accuse + a person + of’ (see line 4). The outcome of her paraphrase was: “. . . this statement harshly accused the US of planning a secret war.” (Refer to Table 4.54 for further details.)

Table 4.54. Subject JYP using *MonoConc* with search term "accus* @ of"

No.	Observation	Conversant	Verbatim
1	VO Typing 'accus* @ of' in M	Researcher	You typed in 'accus* @' and 'of.' Why were you looking for this?
2		JYP	I was looking for expressions containing 'accuse + person + of.'
3		Researcher	What were you looking at?
4		JYP	Accusing her client of corruption.' 'Accuse' is followed by a person, 'of,' and what the person did. So, I thought I could use it. I used to use the <i>Yahoo</i> dictionary to look for examples. This time, I looked for sample sentences in M.
5	VO Typing 'accus* that' in M. Most of examples contain 'accusation.'	JYP	I didn't find any examples with 'accuse + that' without a person in between. So, I used 'accuse + a person + verb-ing' instead of 'accuse + that.' It gave you results with 'accuse + a person + of' before. This time, it gave you 'accusation.' Both of these were used frequently. The expression I wanted to use was 'accuse + a person.'
6	VO Typing 'accused the US of planning a secret war' in W	JYP	Outcome: . . . this statement harshly accused the US of planning a secret war.
OS: It used typically unrestrained language in accusing the United States of secretly planning a war.			
RS: As usual, this statement harshly accused the US of planning a secret war.			

The participant in Table 4.55 typed ‘accus*’ in *MonoConc*. She wanted to be sure that the structure ‘accuse + person + of + fact’ was correct. She saw that this was the most frequently occurring form, so she decided to write ‘accused America’s intention’ in her paraphrase, but later changed it. She saw a preposition, but did not use it at that point. She looked at the results in *MonoConc* again and read over whole sentences. She wanted to know if she could use ‘accuse’ without using ‘of.’ She still wanted to use the formulation ‘accused America’s intention’ and wanted to check if a person had to follow the word ‘accuse.’ She saw that ‘accused’ could be used without ‘of’ and in most of those cases, a person followed, so she changed her phrase from ‘accused America’s intention’ to ‘accused Bush government of their attacking intention’ in her final version (see line 19). She decided this because the most frequent form she found was ‘accuse A of B.’ She did not see a sentence that matched the structure of ‘accuse America’s intention.’ She saw an example in *MonoConc* in which ‘of’ was followed by ‘manipulating,’ so it was an accusation about manipulating, which she related to her writing. In her final paraphrase, she worded the phrase as: “The report accused Bush government of their intention to . . .” Although the phrase contains various grammatical errors, the structure she used, ‘A accused B of C’ is correct.

These excerpts show that all four subjects were able to apply successfully the structure used in ‘accuse’ to their writing. Although some of the sentences containing ‘accuse’ contained grammatical errors, the examples in *MonoConc* enabled each subject to go through a process of elimination to arrive at the correct ‘accuse’ portion of their paraphrases. When subjects were not able to view examples of structures that they thought were correct, this was an indication to them that their guesses were not correct and that they should look for collocation frequency patterns in contexts that were applicable to their content matter.

Table 4.55. Subject YJK using *MonoConc* with search term "accus*"

No.	Observation	Conversant	Verbatim
1	VO Typing 'accus*' in M	Researcher	Why were you looking for 'accuse' in M?
2		YJK	I wrote 'accuse America's intention' instead of 'accuse + person.' I also thought that 'accuse of' was used often. I thought it was possible to write 'accuse + person + of + fact.' I wanted to be sure it was correct.
3		Researcher	What were you looking at?
4		YJK	'Accuse + A + of + B.'
5		Researcher	Was it frequent?
6		YJK	Yes, it occurred most frequently.
7		Researcher	How did you apply it in your writing?
8		YJK	I changed it.
9	AC Pointing at 'accused America's intention' in W	Researcher	Here, you typed 'accused America's intention.'
10		YJK	I changed that.
11		Researcher	You mentioned that you saw a preposition, but you didn't use it here.
12		YJK	I didn't use it at first and moved on. Then, I was going to use 'accuse' again later, and I remembered that I used 'accuse' before. And I wanted to make sure it was correct.
13		Researcher	But you didn't use any prepositions.
14		YJK	I changed it later.
15	VO Looking at the results in M again	Researcher	You were looking at the results of 'accuse' again. Why did you go back?
16	VO Reading sentences by clicking on results	Researcher	And you were reading whole sentences. What were you thinking?
17		YJK	I was wondering how to use 'accuse' without 'of.'
18	VO Mouse pointing at sentences containing 'accuse' without preposition in M	YJK	I was wondering if I could write 'accused America's intention,' and whether 'accuse' must be followed by a person. I found out that 'accuse' could be used without 'of.' In this case, a person followed in most examples, so I changed it later.

Table 4.55. Continued

19	VO Changing 'accused America's intention' to 'accused Bush government of their attacking intention' in W	YJK	I changed it to 'accuse Bush government of' in order to express accusing someone.
20		Researcher	How did you decide this?
		YJK	The most frequent form was 'accuse A of B.' I didn't find a sentence that contained the same structure as 'accuse America's intention.' Most of the examples contain 'accuse A of B,' so I applied it to my sentence. 'Of' is followed by 'accused.' I wasn't sure if I could use 'intention,' which is an abstract noun. It sounded awkward to accuse someone's intentions. I couldn't think of any other words.
		Researcher	How did you find out that an action followed?
	AC Pointing at the results in M	YJK	Here, 'of' is followed by 'manipulating,' so it was an accusation of manipulation.
	AC Pointing at 'accused Giuliani of trying' and 'accused Florio of inflating' in M	YJK	It's about accusing of 'trying' and 'inflating.'
OS: The statement rejected the American formula point by point. Complete nuclear dismantling is a plot to overthrow the North's socialist system . . .			
RS: The report accused Bush government of their intention to “overthrow the North Korea’s socialist system” and repeated the conventional statement . . .			

4.5.2. *MonoConc* leading to wrong grammar

This section focuses on the misinterpretation that occurs in *MonoConc* when nonnative subjects use the examples and incorrectly copy structures into their own writing. The first subject typed ‘as usually’ in *MonoConc* (see Table 4.56). The article used the phrase ‘it is typically,’ so she wanted to replace it with ‘as usually.’ She noticed that ‘as usual’ was an adverbial phrase, so she wanted to find the appropriate place to use it in the sentence. In *MonoConc* she found the example: “The leaves can be toxic not, as

usually thought, because of oxalic acid, but . . .” *MonoConc* has a variety of examples, and sometimes, examples such as these are not exactly clear. If the example sounds strange and is perhaps an unusual expression, it is best for the users to avoid copying this type of structure. This subject did not see that ‘thought’ is a past participle and was used in the passive voice. She wrote: “The radio broadcast blamed United States for planning a war, as usually using offensive languages.” She accompanied ‘as usually’ with the active voice and incorrectly thought that this structure was acceptable.

Table 4.56. Subject JDO using *MonoConc* with search term "as usually"

No.	Observation	Conversant	Verbatim
1	VO Typing 'as usually' in M	Researcher	Why were you looking for 'as usually' in M?
2	AC Pointing at the article	JDO	It says, 'it is typically.' I was going to use 'as usually' to replace 'typically,' but I didn't know where to put it in a sentence. Should I put it in front of a sentence or in the back? If I was going to write 'usually,' I would put it after a verb such as 'use' in this case. But 'as usual' is an adverbial phrase, so I was wondering where the appropriate place in the sentence was.
3	VO Typing 'as usually using' in W	Researcher	You typed 'as usually using.' Why?
4		JDO	In M, it says, 'as usually thought.' It comes before a verb like 'as usually happens.' So, it should be 'as usually using something.'
OS: It used typically unrestrained language in accusing the United States of secretly planning a war.			
RS: The radio broadcast blamed United States for planning a war, as usually, using offensive languages.			

Table 4.57. Subject JHC using *MonoConc* with search term "settle"

No.	Observation	Conversant	Verbatim
1	VO Typing 'America should not settle with any' in W	JHC	I used 'should not settle' instead of 'should not negotiate.' But, I thought 'settle with' was strange, so I looked for it.
2	VO Typing 'settle*' in M	Researcher	You were looking for 'settle*'?
3		JHC	I didn't see 'settle with.' I looked at CF also, but I still didn't find it.
4	VO Typing 'settle' in M	Researcher	Why did you look for 'settle' without *?
5		JHC	I am trying to use the present-tense form. So, I thought I could find similar expressions.
6		Researcher	You will get fewer results with 'settle.'
7		JHC	I was just wondering.
8		Researcher	What were you looking at?
9		JHC	I looked at what came after 'settle.' I was looking for 'with.'
10	VO Looking at CF	JHC	But I didn't see 'with.' So, I decided not to use 'with.'
11		Researcher	What did you decide to do without 'with'?
12		JHC	I'd heard 'settle down' a lot, so I used 'settle down.' But I thought it sounded strange. I then searched for it.
13	VO Typing 'settle @ down' in M	Researcher	'Settle @ down.'
14		JHC	There weren't any results.
15	VO Typing 'settle down' in M	Researcher	'Settle down.'
16		JHC	There were some examples, but they were not what I wanted. I wanted to write a word that meant 'negotiation.'
17	AC Pointing at the results in M	JHC	But I saw 'bird settle down' and 'he is finally ready to settle down.' These were not what I was looking for. So, I decided not to use 'down.' I just wrote 'settle' alone because I saw a lot of examples with 'settle' used as a transitive verb.
18		Researcher	Where did you see that?
19		JHC	I didn't see that many and thought it would be wrong.
20		Researcher	By transitive, you mean a verb with an object?

Table 4.57. Continued

		JHC	Yes. I thought I didn't have to use a preposition because the meaning of this word was the same as 'negotiation.' I saw some examples with 'settle for.' But 'for' wasn't what I was looking for.
21	VO Typing 'settle with' in M	Researcher	You were typing 'settle with' in M.
		JHC	I thought that 'settle' might not be a transitive verb. So, I tried what I could.
22	VO Final results: America should not settle any North Korean's demands that' in W		
OS: The administration has also said it will not provide aid or other benefits to North Korea before it scraps all its nuclear programs and allows rigorous inspections.			
RS: America should not settle any North Korean's demands that refuse rigorous inspection of nuclear program and eventually eliminate its program completely.			

Table 4.57 shows that the participant began with ‘America should not settle with any North Korean’s demands . . .’ in her paraphrase. However, she thought that ‘settle with’ was strange, so she typed ‘settle*’ in *MonoConc* and did not see any examples including ‘settle with.’ She also checked the collocation frequency, but still did not see any examples with ‘settle with.’ She looked for ‘settle’ in *MonoConc* because she wanted to use the present-tense form. Again, she looked for collocation frequencies, but still did not find ‘with’ (lines 9 and 10). After this, she decided to use ‘settle @ down’ as a search phrase because she had heard ‘settle down’ a lot. However, she did not get any results, so she typed ‘settle down’ in *MonoConc*. She did see some examples, but they were not what she wanted. She wanted to see an example with ‘settle’ that meant ‘negotiation.’ Finally, she typed ‘settle with’ in *MonoConc*, but canceled the search quickly when she did not see any examples. She finally decided to write: “America should not settle any North Korean’s demands that refuse rigorous inspection . . .” She

did not think that the preposition was necessary and made an incorrect judgment based on what she saw in the results in *MonoConc*, as shown below.

Table 4.58. Subject JWK using thesaurus with search term “supervise” and *MonoConc* with search term “watch over”

No.	Observation	Conversant	Verbatim
1	VO Typing 'supervise' in T	Researcher	Why did you look for 'supervise' in T?
2		JWK	I thought that the meanings of 'supervise' and 'monitory' were similar. That's why I looked for 'supervise.'
3		Researcher	What did you find out?
4		JWK	'Watch over' seemed to be similar, so I chose it.
5	Typing 'watch over' in M	Researcher	Why did you look for 'watch over' in M?
6		JWK	To see examples to find out how it was used in real life, what came after it, and what to watch over.
7		Researcher	What did you learn?
8		JWK	From the example, I notice that the meaning of 'watch over' was that of 'monitor.' I used it in my writing.
9	VO Typing 'watch over' again in M	JWK	I did this because I canceled the search before, so it didn't give me enough examples.
10	VO Typing 'watch over by news' in W	Researcher	You typed 'watch over by news.' Why did you use 'by'?
11		JWK	I used 'by' to make it a passive sentence. The subject comes last in a passive sentence.
OS: The statement carried by Radio Pyongyang and monitored by news agencies in South Korea came just after a visit to North Korea by China's foreign minister, Li Zhaoxing . . .			
RS: This statement given by Radio Pyongyang and watched over by news agencies in South Korea announced just after coming to North Korea by Li Zhaoxing, China's foreign minister . . .			

In the above episode, the subject typed ‘supervise’ in the thesaurus because she thought the meanings of ‘supervise’ and ‘monitory’ were similar. She discovered that ‘watch over’ was similar, so she chose to type this phrase into *MonoConc*. She wanted to

see how it was used in real life, what word or words followed it, and what should be ‘watched over.’ Because she saw that ‘watch over’ had the same meaning of ‘monitor’ in her *MonoConc* example, she decided to use it in her writing. However, she had to type ‘watch over’ again in *MonoConc* because she had canceled her previous search, and had not seen enough examples. Her final choice of wording was: “This statement given by Radio Pyongyang and watched over by news agencies . . .” which is actually inappropriate wording in this sentence. Although ‘watch over’ and ‘monitor’ are synonyms, they are not used in the same way by native English speakers. The term ‘watch over’ indicates the action of looking after a young child, while ‘monitor’ is used in a more professional sense. In her sentence, the news agencies ‘watched over’ the statement given by Radio Pyongyang, which sounds strange to the native ear. This excerpt is another example of how nonnative speakers may misinterpret an example given in *MonoConc*, although it is very close to the phrase that they are looking for. See Table 4.58 for further details.

These three examples demonstrate that simply using *MonoConc* as a reference tool does not guarantee that all writing is going to be perfect. At times, subjects do not rely on their ‘gut instinct’ and rely only on what is indicated in *MonoConc*. Corpora found in *MonoConc* and other concordancing programs do not constitute a complete listing of all possible forms. Although many examples are offered, users must not assume that if their entry word or phrase is not found that it is not correct. Additionally, users may not always thoroughly search the examples in order to find the context and or structure they are looking for. Another reason for unsuccessful *MonoConc* searches is that the search words or phrases are inadequate. They may be misspelled, need to be more generalized or more specific, or add or take away search characters. Nonnative users may have an especially difficult time with some examples that show unusual forms or those that are rarely used by native speakers and may misinterpret them simply because they are too complex.

4.5.3. Using phrases as search terms

Among the many convenient features that *MonoConc* offers, the function that allows users to use a phrase as a search term was very effective in terms of saving time and providing clarification. Also, when a phrase is used as a search term, *MonoConc* shows examples of very specific cases of structure and context. This section focuses on how subjects used phrases for searching and the effects of this strategy on their writing.

The phrase that the first subject, JYP, searched for (see Table 4.59) was ‘conflict over.’ At the beginning she had written ‘conflict about,’ but thought about using ‘conflict over’ or ‘conflict on.’ She wanted to see which phrase was used most frequently. She also looked for ‘conflict on’ in *MonoConc*, but did not obtain results. Her third search phrase was ‘conflict about,’ but it was not used very frequently, so she decided to use ‘conflict over’ because it provided more sample sentences. Her search for the expression ‘conflict over’ was quick and she decided on it by the frequency she saw in the *MonoConc* results.

Table 4.59. Subject JYP using *MonoConc* with search term "conflict over"

No.	Observation	Conversant	Verbatim
1	VO Typing 'conflict over' in M VO Typing 'conflict on' in M VO Typing 'conflict about' in M	JYP	At first, I wrote 'conflict about.' Then, I thought about using 'conflict over.' I also thought about 'conflict on.' So, I wanted to find out which one was more frequently used among those three. I looked for 'conflict on,' but it didn't give me any results. I used 'conflict about.' But it wasn't used frequently. So, I decided to use 'conflict over' because there were more sample sentences.
4		Researcher	You wrote 'conflict over' before looking for it.
5		JYP	Yes. I wanted to find out if this expression was correct, and if 'about' was used more frequently.
OS: Bush administration officials have repeatedly stated that they will not sign any agreement with North Korea that does not use that wording.			
RS: The US and North Korea have the tense conflict over the use of specific wording in the agreement.			

Another example from this same subject was her search for the phrase ‘made it clear that’ (Table 4.60) in *MonoConc*. She said that she had created this expression while paraphrasing and needed clarification to know if it was correct. She saw the examples in *MonoConc* and the context matched that of her phrase. The structure also matched her structure, so she decided to use her own phrase and did so correctly.

Table 4.60. Subject JYP using *MonoConc* with search term "made it clear that"

No.	Observation	Conversant	Verbatim
1	VO Typing 'made it clear that' in M	Researcher	You were looking for 'made it clear' in M.
2		JYP	I made up this expression when paraphrasing. I wrote this expression in W and wanted to find out if this was correct. I was wondering in what kind of context it was used.
3		Researcher	What were you looking at?
4		JYP	If the context of the sentence I was going to write matches with that of the sample sentences in M.
5	VO Reading the results in M	Researcher	What were you reading?
6		JYP	Here, it says 'she made it clear that.' 'It' and 'that' were used in the same way I wanted to use them. Also, there were many examples used in the newspaper. So, I thought I could use it that way.
OS: Bush administration officials have repeatedly stated that they will not sign any agreement with North Korea that does not use that wording.			
RS: . . . the US made it clear that they will sign only when the contract includes that wording.			

The second subject (Table 4.61) typed ‘negotiation talks’ in her writing first, but then used ‘negotiating table’ as her search phrase in *MonoConc*. She said that she had used ‘negotiating talks’ before, and that she wanted to change the structure. She wanted to see if it was correct and if it was used in news articles. She decided to use it and changed ‘negotiation talks’ to ‘negotiation table.’ Her final product would have been correct, had she remembered to change ‘negotiation’ to ‘negotiating.’ This subject’s

search phrase was for the purpose of a quick check on, or clarification of, what she already thought was correct.

Table 4.61. Subject YJK using *MonoConc* with search term "negotiating table"

No.	Observation	Conversant	Verbatim
1	VO Typing 'negotiation talks' in W. Typing 'negotiating table' in M	Researcher	Why were you looking for 'negotiating table in M?
2		YJK	I used 'negotiating talks' before, and I wanted to change the structure. I remembered 'table,' so I used 'negotiating table.' And I was wondering if it was correct. I wanted to see if this kind of expression was used in news articles. I used 'negotiation talks' before, but I didn't want to use the same expression repeatedly. I changed it to 'table,' but I wasn't sure if this expression could be used in news articles. So, I looked for it.
3		Researcher	What did you decide?
4		YJK	I thought I could use it because it was used here.
5	VO Changing 'negotiation talks' to 'negotiation table' in W	YJK	Didn't I change 'negotiation table'?
6		Researcher	You wrote 'negotiation table' not 'negotiating table.
7		YJK	I thought I wrote 'negotiating table.' I was looking for 'negotiating table,' so I thought I used the same phrase in W. I didn't check it.
OS: Bush administration officials have said that North Korean negotiators discussed accepting the American terms as a common negotiating goal in the last round of talks . . .			
RS: Bush government was expecting tangible progress in the final negotiation table, keeping its position about disposing holistic nuclear program in North Korea.			

As shown in the results, *MonoConc* can be used effectively for sentence-level composition. Search phrases are useful for users that already know a phrase, or think that they know one, but want to check for quick clarification. Collocation frequency, or lack

thereof, is also a key factor in determining which phrase the users will ultimately choose. Users are not misled by *MonoConc* when they are able to relate the contexts of the examples to the context of their paraphrases. On the other hand, the users are misled when they do not understand the context in which a given example is used.

4.6. What are learners' attitudes toward using a corpus as a reference tool?

The following section is a compilation of students' opinions and reactions toward using *MonoConc*. Research data for this question was obtained from the survey. The survey contains questions that focus on effectiveness, helpfulness, and ease of using *MonoConc* for the subjects' writing (refer to Appendix G for the research questions).

The surveys were measured by the mean and rate of standard deviation (see Figure 4.8 and Table 4.62). All participants agreed in their surveys that the practice session previous to the actual research was helpful and that the searching technique was easy to learn. Almost everyone agreed that the corpus was more helpful than a dictionary for English writing. Half of the participants agreed that using the corpus was helpful for finding out the meaning of vocabulary, while the other half did not agree as strongly. The majority did agree, however, that the corpus was helpful for finding the correct usage of vocabulary and phrases. All participants agreed that the corpus was helpful for finding out grammatical usage, improving English writing skills, and increasing their overall accuracy in the grammar of their English writing. The majority of students did not have much difficulty using the corpus due to time and effort spent on data analysis, or unfamiliar vocabulary in concordancing output. Participants did not experience much difficulty with the corpus due to finding cut-off sentences in concordancing output or having too many sentences produced by the concordancing program. On the other hand, the participants marked on their surveys that they had some difficulty using the corpus due to obtaining a limited number of sentences in the concordancing output. Most

participants did not have undue difficulty analyzing concordancing and collocate output. Participants had minimal difficulty in performing search techniques and the texts presented in the corpus were understood. Everyone agreed that they now understood the purpose of using the corpus for writing and want to use it the next time they have a writing assignment. All participants also agreed that they searched for help in the corpus when they encountered problems with their writing and usually found the information for which they were searching. A minority of students stated that they used the corpus for purposes other than academic writing, and almost all respondents agreed somewhat strongly that they would use the corpus for English writing in the future. All participants agreed that their confidence in writing English has increased thanks to the corpus and they believe they would have had more accurate papers and writing in general had they used the corpus earlier. Everyone also agreed that the corpus was a very useful resource for his or her writing. They all agreed that corpus was more useful for writing than for reading in English. The entire group was in agreement that other international students and ESL students should be introduced to corpus usage and that they would recommend using corpora in ESL writing courses in the future. All of the participants were originally from Korea and believed that corpus use should be taught in English classes in Korea. They all recommended the corpus to other international students at their university or elsewhere.

Table 4.62. Survey questions and results

No.	Question	Mean*	SD
1	The searching technique was easy to learn.	1.4	0.52
2	The practice session last week was helpful for learning the techniques.	1.4	0.52
3	The corpus is more helpful than a dictionary for my English writing.	2.7	1.16
4	Using the corpus is helpful for finding out the meaning of vocabulary.	3.4	1.65
5	Using the corpus is helpful for finding out the usage of vocabulary.	1.4	0.52
6	Using the corpus is helpful for finding out the usage of phrases.	1.5	0.71
7	Using the corpus is helpful for finding out grammar.	1.8	0.63
8	Using the corpus improved my English writing skills.	2.31	1.64
9	Using the corpus increased the grammatical accuracy of my writing in English.	2.33	1.35
10	I have some difficulty in using the corpus due to limited access to computers.	4.6	1.43
11	I have some difficulty in using the corpus due to time and effort spent on analyzing the data.	5	1.5
12	I have some difficulty in using the corpus due to unfamiliar vocabulary in concordancing output.	4.78	1.3
13	I have some difficulty in using the corpus due to cut-off sentences in concordancing output.	4.56	1.33
14	I have some difficulty in using the corpus due to too many sentences in concordancing output.	4.44	1.01
15	I have some difficulty in using the corpus due to the limited number of sentences in concordancing output.	5	0.71
16	I have some difficulty in analyzing concordancing output.	4.44	1.42
17	I have some difficulty in analyzing collocation frequencies.	4.44	1.24
18	I have some difficulty in performing the search techniques.	5.33	1
19	The real texts in the corpus are too difficult to understand.	5.22	0.67
20	I understand the purpose of using the corpus for my writing.	1.3	0.48
21	I want to use the corpus for writing in English again.	1.6	0.97
22	When I have problems in writing in English, I search for help in the corpus.	1.7	0.67
23	When I search for information in the corpus, I usually get the information that I need.	2.3	0.82
24	I use the corpus when writing for purposes other than academic writing.	3.13	1.36
25	As I have learned more about the corpus, I have come to like it more.	2.8	1.4
26	I will use the corpus for writing English in the future.	1.7	0.67

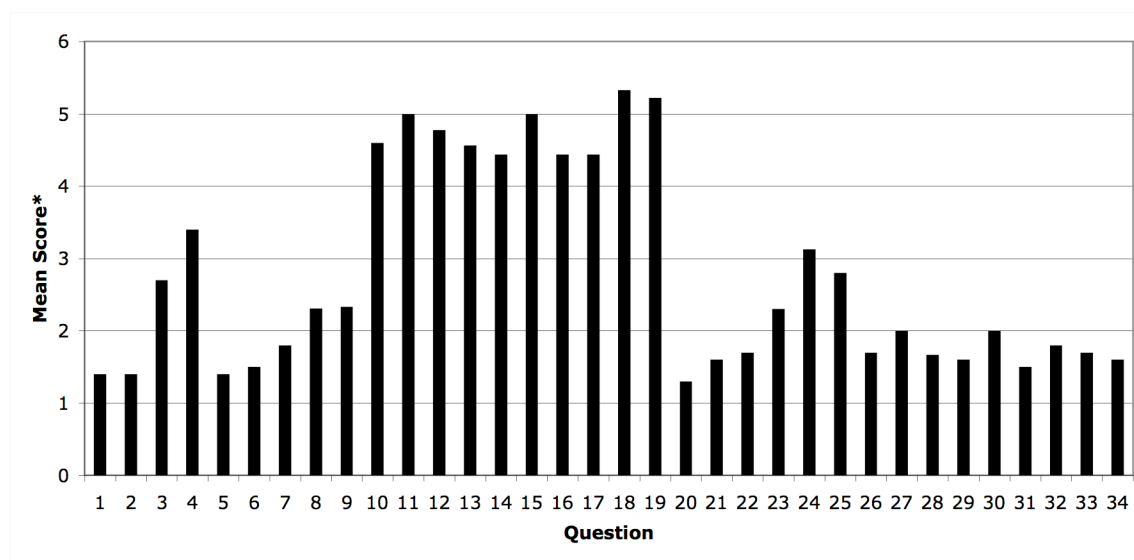
Table 4.62. Continued

27	Learning about the corpus has increased my confidence about writing in English.	2	1.05
28	If I had used the corpus earlier, I would have had more accurate papers or any types of writing.	1.67	1
29	Overall, the corpus is a very useful resource for my English writing.	1.6	0.52
30	Corpus use is more helpful for writing than for reading in English.	2	0.82
31	The corpus should be introduced to other international students and ESL students.	1.5	0.53
32	I recommend using the corpus in the ESL writing courses in the future.	1.8	0.63
33	Corpus use should be taught in English classes in Korea.	1.7	0.48
34	I will recommend the corpus to other international students at the University of Iowa or elsewhere.	1.6	0.52

SD: Standard Deviation

*1:strongly agree, 6: strongly disagree, 0: no opinion

Figure 4.8. Survey results



*1:strongly agree, 6: strongly disagree, 0: no opinion

CHAPTER V

DISCUSSION

This chapter discusses the results of this study in terms of the purposes of using reference tools, the information looked for, the learners' strategies, the effects of using *MonoConc* for accurate writing, and the attitudes of the subjects toward using *MonoConc*. The implications of corpus use for language teaching and future research are also presented in this chapter.

5.1. Purposes of using a concordancing program

This section discusses the purposes of using reference tools and the kinds of information looked for. Purposes for using *MonoConc* include collocations, definitions, context, and parts of speech. The subjects looked for a variety of information in *MonoConc*, including prepositions, authentic samples, and the context in which the search terms were used. Reasons for using dictionaries include definitions, parts of speech, and sample sentences. The thesauri were used to look for synonyms.

Collocation was the most common reason learners used *MonoConc*. When looking for collocates of words, the most common combination was verb + preposition. Prepositions often gave the subjects difficulty even though they were advanced learners, which accounts for the large number of searches conducted for verb + preposition combinations. Concordancing programs made it possible for the subjects to have direct access to verb phrase combinations. Not only were they able to find verb phrases, but they were also able to compare findings, in order to find the specific combination that suited their needs. The subjects in the present study preferred seeing a variety of prepositions to choose from in order to make the best choice. At times, seeing one example was enough information for the subjects, but more often than not, the subjects, as a whole, made choices based on the frequency of the collocations in the examples

provided. Popular resources for both students and teachers are grammar books, which often provide a list of verb + preposition collocations, but findings in *MonoConc* are much more exhaustive and are authentic.

According to Aston (1996), learners stated that when comparing a concordancing program to ‘conventional reference instruments,’ they were more successful in using the concordancing program with collocates for that particular situation. These results echo Yoon’s study (2005), in which the subjects also used the concordancing program to look for collocations such as prepositional phrases, phrasal verbs, and modifiers. Yoon presented an overall impression of how her subjects used the concordancing program, but the data gathered (user logs, retrospective interviews, classroom observation, and a survey) did not document detailed interactions of concordancing program use.

In the present study, the results show detailed interactions between the subjects and the concordancing program. Using *MonoConc* for writing is a complicated procedure, involving looking for possible collocations, based on what was found in dictionaries, previous *MonoConc* searches, or previous knowledge. The screen recordings showed the subjects’ actual use of the concordancing program, with a stimulated recall session that immediately followed. These two different kinds of data demonstrated what the subjects saw and thought and how they decided to apply the findings to their writing. Subjects applied their findings to sentences and contexts that they were trying to produce. Such multilevel thought processes and interactions are described in this study.

When subjects in this study looked for collocations, one of the main reasons was to find the preposition following the verb. When they looked for the preposition, they entered a verb that they believed accompanied the preposition. After using the verb as a search word, they found some possible prepositions or structures and narrowed their searches by entering several words at a time. They analyzed not only examples, but also collocation frequencies. Most of the time, they concentrated on looking for frequencies

of collocations and preferred to use frequently occurring collocations in their writing. However, sometimes they had a theory of the type of collocation that could be used with a word, and looked for specific combinations. When they saw these combinations, they believed that it confirmed their theory. If they did not see the combination they were looking for, they would choose to look for different terms among the collocations they found.

Yoon observed in her study that the collocation frequency data were not used very much by the subjects. This is primarily because the concordancing program took a longer time to produce frequency data and because the students wanted to look at how words were used in sentences. In contrast, the subjects in the present study used collocation frequency data more often, possibly because the concordancing program provided collocation frequency data immediately and because the format in which *MonoConc* displayed collocation frequency was user-friendly, so that it made sense to the users. The concordancing program used in Yoon's study only provided a list of collocations and their occurrences and was not as user-friendly. Another factor may have been that the subjects in the present study were more advanced. When they saw lists of collocations, they were more familiar with collocations than those in Yoon's study.

Secondly, a concordancing program was utilized for the purpose of finding definitions. Subjects looked at relationships between search words and surrounding words in the examples provided in order to reach their conclusions. The sampling *MonoConc* provided also helped them to see if the words or expressions they were considering using actually existed in real life situations. Subjects not understanding the appropriate context in which words are used could rely on *MonoConc*, as they deduced the meaning from the various contexts provided. Since ESL learners may lack the social experiences of hearing new vocabulary in various contexts, implementing a concordancing program gives an insider's perspective on subtle nuances.

Verification occurred when subjects conducted searches to see whether their writing was correct or not, and was another reason for them to consult a concordancing program. These advanced ESL subjects often came up with expressions that were not exactly native-like, so they consulted *MonoConc* to see if their theories actually could be exemplified in authentic native speakers' speech and rephrased their wording if necessary. Similarly, the subjects decided to use phrases that they previously thought had sounded awkward after seeing expressions with the phrases in *MonoConc*. This is a challenge often faced by ESL learners, as they will repeatedly commit the same errors due to interference of their primary language's grammar, contexts, tonalities, pronunciation, and the like.

The subjects believed that if what they were looking for occurred frequently in examples given by a concordancing program, then what they intended to use was correct. However, when they had a specific structure in mind, regardless of the number of occurrences, the subjects copied it into their writing. They trusted in *MonoConc* enough to believe that one appearance occurring in authentic texts was enough to prove that the intended structure was correct. Understanding context was another use for *MonoConc* in this study. The surrounding words were what the subjects generally looked for as well as the application to that particular situation. By looking at the context, the subjects were able to not only discern the meanings, but also the situations in which this specific word or phrase was used. Although there are both supporters (Frantzen, 2003; Stevens, 1991b) and critics (Haynes, 1993; Schatz & Baldwin, 1986) of how well context helps learners infer the meanings of words, this is an issue well known to ESL learners. Context is rich in nuances not familiar to those who have grown up in a different culture, so implementing a concordancing program serves as a quick research tool that assists ESL learners in deciphering not only the meanings of words, but also the complex contexts in which they are used.

The subjects sometimes used a concordancing program and dictionaries in order to find the parts of speech by analyzing phrasal structures surrounding the search words. When they went to the dictionary to find the part of speech and if the word had only one form, they could quickly understand its function in a sentence. However, if words had a noun form and a verb form, the subjects went to *MonoConc* and found the correct form by looking at the surrounding words. This analysis requires understanding of syntactic structures of phrases and sentences. When learners are able to carry out this type of analysis, *MonoConc* can be a quick reference tool, not only for collocations, but also for syntactical understanding. This can also be found by analyzing the collocation frequencies for words coming before and after the search word.

Dictionaries and thesauri also played a role as reference tools in this study. The subjects used the thesaurus mostly to search for synonyms or word changes. Whereas ESL learners may function well in everyday speaking situations, they may lack the normal range of choices in vocabulary that native speakers possess. This deficiency can be particularly troubling when learners are given the task of writing, where the overuse of words becomes visually apparent. When dictionaries were used, the purpose was mostly to find the definition and synonyms. Subjects also used dictionaries to find examples of certain usages of the search word. When subjects knew the definition of the word, they used dictionaries to make sure that their theory of the definition was correct. Because the task was paraphrasing, they looked for synonyms to change the words used in the original text to a similar word. Then, they analyzed usages of the word in sample sentences provided by the dictionaries. At this point, the subjects decided if they needed to use *MonoConc* for more information about collocations and structures.

The English-Korean dictionary was utilized mostly for definitions and for word changes or synonyms. The Korean-English dictionary was also used, but the majority of the searches (23 times) were for translations, while only 12 searches performed were for synonyms or word changes. The least used tool was the English-English dictionary,

which, for the most part, was used for finding definitions, synonyms, or word changes. This finding somewhat contradicts Baxter's claim (1980) that paraphrasing encourages monolingual dictionary use. However, the subjects in the present study did not take advantage of the definitions provided in the monolingual English dictionaries. This could be because the definitions in the English-Korean dictionary were easier to understand or quicker to read. Additionally, by using the English-Korean dictionary, subjects may have gotten an accurate definition for the search word. In other words, they may have been able to grasp subtle and multiple meanings of the word more easily than when using the English-English dictionary.

When using a thesaurus, the most common search terms subjects used were verbs. This is because they tended to change the main verbs in the original text when paraphrasing. In the thesaurus, they would find the verb that they thought was most appropriate for the context and decided whether it would then be necessary to use a dictionary or *MonoConc*, or simply to use the verb they had already found.

By analyzing the examples when using a concordancing program, subjects come up with their own definitions, collocations, and theories of syntactic structures. Since a concordancing program does not give any direct information about what to look for, one of the subjects' tasks was to choose and adjust the search terms as they saw fit. The extent to which a concordancing program helps the subjects depends on how they perceive information and extract what they needed from what they see.

5.2. Learner strategies for using a concordancing program

The most common macro strategy was to combine tools (47 times, 48.0%), while the second most common was to use a specific search word (13 times, 13.3%). Six times (6.1%), the subjects performed search word changes or used special characters. Subjects who used more than one tool for a search or performed multiple searches were more successful in finding what they were looking for.

Dictionaries and corpora individually benefit learners, but when used in conjunction, they complement each other. While corpora do not directly give definitions, they do provide many authentic examples, which enable learners to come to their own conclusions about search words or phrases. Yoon found that her subjects referred to the dictionary first for definitions and then the corpora as a “linguistic checker” for usage. However, she only showed a general picture of how the subjects combined dictionaries and the concordancing program from retrospective interviews about students’ opinions of the reference tools.

The present study shows more detailed information about what dictionaries provide for different purposes and how *MonoConc* complements them. The subjects entered search words first into the English-Korean dictionary to see what words could follow them. After seeing that a noun could follow verbs, they used *MonoConc* to check for more examples. Finally, they were able to make a decision by seeing that nouns come after verbs. The availability of correct examples of collocations enabled the subjects to make informed choices. The subjects in this study depended heavily on *MonoConc*, as they saw that the reliability of the results available to them through using special characters, different entry combinations, and wide array of corpora in the concordancing program.

The subjects used a combination of strategies in order to find their information. For example, the subjects often typed a word into the dictionaries and thesauri, and saw the results. Next, they copied the results into *MonoConc* to check for context and part of speech (noun or verb). They also confirmed the words that came before and after the search word, and were able to make a decision. Consulting both the thesauri and *MonoConc* clarified the search information—they were able to find synonyms, context, parts of speech, and collocation words.

The combination of thesauri and concordancing program occurred when the subjects needed to change a word (usually verbs), but did not know collocations or how

to use a word in a sentence. In this case, they went to a thesaurus and chose the most appropriate word. When they had what they thought was the most appropriate word, they decided whether or not to go back to the dictionary. If they were not completely sure, they went back to the dictionary and found more information about the word, such as definitions and sample sentences. On the other hand, if they knew the meaning of the word, they only used *MonoConc*. In this way, *MonoConc* complemented the thesaurus by providing quick access to collocation frequency, which cannot be found in a thesaurus.

In conjunction with the dictionaries, a corpus could be a useful linguistic resource to refer to for immediate help. When the subjects used a combination of dictionaries and *MonoConc*, they used *MonoConc* as a quick reference for confirmation. When they did not know the meaning of a word or how to use it in a sentence, they first looked in the dictionaries for definitions and synonyms to use in their writing. Next, they went to *MonoConc* to see how to use the words in the sentences and what the collocations were. When they knew the meaning, but did not know the collocations or how to use it in a sentence, they did not use the dictionary and went straight to *MonoConc* for sample sentences and collocation frequencies.

A concordancing program was used as an instrument that confirmed the first finding that users see in the dictionary. Subjects in this study searched for words in the English-Korean dictionary, then typed the search words again in *MonoConc* and noticed that there were patterns in both the English-Korean dictionary and *MonoConc*. Seeing grammatical patterns in various sources helped confirm the correct answers to those doubting that the answers in the dictionaries were correct. This method of double-checking is much more reliable, as the definitions and examples displayed in dictionaries do not always sound authentic.

Some subjects also used the Korean-English dictionary in conjunction with *MonoConc*. In this study, subjects used them to understand sentence structures. First

they looked for English equivalents of the Korean word. After finding the English translation, the words were entered into *MonoConc* to see how they were used in sentences. Again, *MonoConc*'s authentic display of wording and structure helped confirm the best choices.

Subjects also looked in the Korean-English dictionary to find slang terms. When they could not find satisfying answers, they chose other words in the dictionary. Using the translation, they would then enter it into *MonoConc*. After identifying the proper nuance for a word they entered in the Korean-English dictionary, they were able to make decisions. Slang terms can be particularly troubling, so using a concordancing program can be helpful, as it provides authentic examples, whereas examples in dictionaries sound more artificial. Additionally, as language is continually changing, a concordancing program is the best choice for finding up-to-date terminology.

A combination of a concordancing program and dictionaries was also used when subjects tried to see how a word functioned in a sentence or to test their theories of grammatical structures. If they did not think that the meaning of the search word was different after analyzing the samples, they used a dictionary or thesaurus to find the most appropriate word. If they did not like the collocates with the search words, they referred to the thesaurus to look for sentences, and then back to *MonoConc* to find collocations. If they thought that the nuance of the word that they wanted to use was not what they thought it to be after analyzing the examples, they would check in the thesaurus for synonyms. Subjects did not completely rely on the English-Korean dictionaries because the definitions sometimes sounded awkward in Korean. This is why they referred to a concordancing program first in order to test their theories of meaning and collocations, followed by the English-Korean dictionary.

Another strategy the subjects adopted was entering special search characters for effective searches such as the asterisk (*) and the at (@) symbol. When the subjects wanted to find examples including all possible forms of the search word in *MonoConc*,

they used the asterisk. By using the asterisk, they were able to find the varying forms of the root word they were searching for. In this way, the subjects were not only able to know the noun form of a word, such as a ‘plant,’ but they were also made aware that a verb form (‘to plant’) also exists. The subjects also used at (@) symbols in order to search for examples including the search word with the specified collocate. By using this option, they were able to pinpoint exactly what they were searching for. The combination of using special search characters and comparing structures helped the subjects make the best decision. Using special characters opens up more options that may not appear otherwise, based on the limitations of the search words or phrases themselves.

Another useful strategy is when the subjects performed multiple searches in *MonoConc*. When they were looking for a specific search word and the first inquiry did not render what they were looking for, they changed their search to find a similar structure. Users of *MonoConc* may not always find the answer they are looking for on their first search. When this happened, the subjects changed the search word. When subjects typed a search word in the concordancing program and could not find anything appropriate, they continued changing the words they wanted to use, although they did not always find a word that they liked. This occurred when the search term was too broad and they therefore had a difficult time coming up with a grammar rule that could apply to their sentence. In this case, they could add more words to the search in order to narrow it down. For example, using prepositions with a verb as a search phrase could lead to more specific examples. Alternatively, subjects also began with a very narrow search and when results were not displayed or the results were too few, they could broaden their search. For example, when they searched using a verb + preposition and did not get the results they wanted, they could broaden the search by dropping the preposition in the *MonoConc* entry. The more creative learners can be when using a concordancing program, the

better. When learners have difficulty coming up with alternative search words, they can brainstorm, listing any words that come to mind for their specific topic.

An additional strategy subjects used was to make a decision based on the absence of examples in *MonoConc*. When an example that they were looking for did not appear, they assumed that the expression might not be grammatically correct. As a result, they decided not to use the expression. However, learners should be aware that data listed in a concordancing program, although thorough, may not constitute an exhaustive list of all possibilities.

The results showed that the subjects implemented a number of strategies that were explored during the course of the study. By combining reference tools, the subjects were able to manipulate them and find the results that they were looking for. In order to do this, they had to be able to note the advantages of each tool and understand what each tool provides. For this to be effective, subjects needed time and practice to get accustomed to using the concordancing program in conjunction with other reference tools.

5.3. Corpus use and accuracy of writing

For language learners, a corpus can be a very valuable resource that helps them become more proficient in their L2 writing. A concordancing program enables users to see multiple examples of everyday language use. By using the concordancing program, learners can see words that are used most frequently, their patterns, and collocations. It is common that users of concordancing programs believe that the most grammatical expressions or word combinations are the ones most commonly cited items in a concordancing program. Learners using the concordancing program take more responsibility for their language learning, as they become researchers in their own right. They gain confidence as L2 writers as they have direct access to linguistic resources. L2 writers become more independent and able to solve their own writing and linguistic

problems as they become more aware through the use of authentic texts. In this study, the subjects found the corpora to be more useful for sentence-level composition and revision, rather than global content, discourse, or organizational levels.

The results show that use of reference tools leads to an improvement in the accuracy of writing. *MonoConc* played an important role in defining the structure and context of English phrases and sentences. Compared to *MonoConc* use, the use of other tools was significantly lower. Although the ratio of correct outcomes was very high, one cannot say that a specific tool contributed more than others to the accuracy of the writing outcomes. The following are possible explanations of why one cannot conclude that a specific tool contributed to the accuracy of writing more than other reference tools. Throughout the study, the subjects frequently used combinations of reference tools. The subjects had different purposes for selecting one tool over another. Sometimes the tools provided the same types of information, so that one tool could not be credited with helping more than other tools. Sometimes subjects used previous experiences or prior knowledge that they had pertaining to the writing. At other times, subjects used the tools and found the correct answer, but later revised it and made a mistake. In this case, it is difficult to define whether the information they saw caused this mistake or if there were any other factors involved. In conclusion, a concordancing program and other reference tools may help learners produce accurate writing by complementing one another.

Although there are many advantages to using a concordancing program, L2 writers can sometimes make incorrect decisions based on the examples provided. When encountering examples that appear unusual, they may be ill advised to copy this expression to their writing, especially if they are not advanced-level language learners. Before using the structure they find in a concordancing program, they should verify that the context is the same as the one they are using.

One of the features that the online dictionary does not support is allowing L2 writers to search for a combination of words. A concordancing program, on the other

hand, allows L2 writers to search for phrases, as well as individual words. By allowing the phrasal search, the subjects in this study benefited from making their phrases sound more native-like. By performing phrasal searches, they quickly accessed the patterns they needed. If the subjects used dictionaries, then they had to choose one of the words in the phrase, search for it, and find the definition matching their inquiry. On the other hand, when they used *MonoConc*, they could use any combination of words or special characters, resulting in more examples than any dictionary alone could provide.

MonoConc provided quick access, the subjects saved time, and could view a number of authentic examples through which they came to a better understanding of authentic English grammar and sentence structures. Oftentimes, if they found examples that exactly matched their sentence structure, they copied what they found into their writing.

Yoon (2005) also describes the advantage of using a concordancing program for writing. She conducted one think-aloud session to find out how the subjects came up with phrasal or sentence structures by analyzing corpus data. She found that the concordancing program helped the subjects identify unique features of words and categorized them. However, her data did not allow her to describe the complexity of the relationship between the concordancing program and the subjects' writing. That is, her data did not display what was going on in the subjects' minds in dealing with various problems they faced when writing and using the concordancing program. The present study shows various types of problems the subjects faced when writing, what they looked for, what they found, and how and why they applied the findings to their writing. Dealing with the information found in dictionaries and a concordancing program and applying it to writing is a very complicated procedure. Subjects not only copied the results from the reference tools, but also tried to fit the structure and context of the information into their writing. When they did not find anything appropriate for the sentences they were trying to write, they referred to *MonoConc* to find other collocations that were more suitable for their writing. Sometimes they did not use what they found in

reference tools and came up with totally different words, phrases, or sentence structures when necessary.

Overall, the subjects received a tremendous amount of help from the concordancing program for their writing. Based on an analysis of the data from the stimulated recall sessions, the concordancing program showed collocations and contexts that the subjects referred to in order to find meanings of search words. By looking at the surrounding words, they also identified contextual information through authentic examples. Another advantage of using a concordancing program was that it provided quick access not only to the examples, but also to the listing and frequency of collocations that the subjects used in order to find the most common combinations of the search word and collocations. This may reinforce learners' awareness of collocations and their use. A concordancing program can also be used in conjunction with dictionaries and thesauri because it provides an abundance of examples that other reference tools may not. The subjects took advantage of the functionality of the concordancing program, as it allowed phrasal searches, so that they could see very specific examples according to the sentence that they were trying to construct. This could potentially increase the accuracy of subjects' writing and confidence in correcting their own writing if they strategize and are cautious in their decision-making.

5.4. Learners' attitudes toward using a corpus as a reference tool

As expected in this research, the reaction of the subjects with high English proficiency to using the corpus and concordancing program was positive. They agreed that it was more helpful for writing than depending on a dictionary as their sole reference tool. This could be because subjects with high proficiency may know the meanings of many words, but may have a difficult time actually choosing the best one when writing. By providing actual samples of words and phrases, the concordancing program helped

learners gain confidence in their English writing. This echoes the findings in Yoon's study (2005): writers became more independent and confident when there was increased availability to linguistic resources; by using authentic texts, students become more aware of stylistic issues and are therefore more independent in solving their own linguistic and writing questions.

The subjects in Yoon's study (2005) complained that the procedures required to access the corpus, look for data, and produce results were too time consuming. The time they spent analyzing the data was another complaint. Their attitudes gradually changed from skepticism to agreement and enthusiasm in using the concordancing program as they became more familiar with the technology. In Yoon's study, language proficiency played an important role in the subjects' abilities to make sense of the data.

Although some of the subjects in Yoon's study (2005), complained about the amount of time spent on searching for data and analyzing it, none of the subjects in this study experienced a difficult time in dealing with a large amount of concordancing output. Only one of the subjects in the present study mentioned that the effort it took to analyze the data while writing a final paper for his own class was very time consuming. This discouraged him from using *MonoConc* somewhat, because he did not have time to search for the quantity of words he deemed necessary. Yoon mentioned that not only efficiency, but also language proficiency affected the attitudes of the subjects in her study. However, this was not a concern for subjects in the present study, nor did the subjects have a difficult time in dealing with cut-off sentences in the concordancing output. The subjects did, however, mention in the surveys that when a limited number of sentences were displayed in the concordancing program, it was difficult to apply this information to their writing. This is somewhat contrary to the findings in the Yoon and Hirvela study (2004), where ESL students used a corpus as a reference tool for writing. The results from their survey show that the students complained of having too many

examples to analyze. This could have been because these learners had a lower level of English proficiency.

By contrast, the learners in this study had a higher level of English proficiency and were thus able to deal with a larger quantity of information and could determine which pieces of information were useful and those that were not. They were able to benefit from looking at the large number of examples and the patterns found in them. They also had more chances to find the context that they were looking for because they had a higher ability to understand the variety of contexts given.

5.5. Implications for corpus use in language teaching

Corpora and concordancing programs can be motivating and rewarding for language learners and teachers. Corpora provide contextualized examples that learners can use for hard-to-answer lexical questions. As the results show, the subjects often combined dictionaries and thesauri with a concordancing program. By combining these reference tools, the subjects were able to understand the context and find collocations of the search terms in the authentic examples. The convenience of having authentic language or 'native-like language', at learners' fingertips can dramatically alter the learning experience both in and out of the classroom. By using concordancing programs, language learners can benefit from learning inductively and taking responsibility for their own learning. However, concordancing programs have not been and presently are not widely used by language learners and teachers because they have not been predisposed to these ideas. Teachers and students may choose traditional instructional methods because they are more comfortable with them. Students in this study were using a concordancing program for the first time. Although it had never been a part of their language learning, they quickly realized that it could be very valuable.

Students need some time to become accustomed to using the concordancing program and develop strategies for using it. During this study, the subjects only had one

week between the tutorial and recall session. As a result, they did not have very much time to become acquainted with using the program and implementing searches with the special characters, nor did they have a large amount of time to apply this to their writing. One month after the study, there was one student who realized how valuable the program was and requested information about how to access the corpus and the program for academic writing purposes. This suggests that it is important to include carefully designed activities that can help develop students' researching skills, making the case for all of the functions the concordancing program can provide and how useful this information can be to their writing. Activities should also include suggestions as to how the concordancing program can be used in conjunction with dictionaries and thesauri, as well as the types of information that can be gained from the program in comparison with other reference tools.

The ideal concordancing program for students must be quick and responsive. It needs to access a corpus quickly, so that valuable time is not wasted while accessing the database. During the course of this study, the subjects mentioned that in the past they were not usually motivated to use dictionaries and thesauri because using those reference tools slowed down their writing process. They were concerned that this new concordancing tool may also take too much time. Some concordancing programs that work as stand-alone programs may process the database quickly and show numerous examples instantaneously as learners perform corpora searches. In this case, generally the faster the computer is, the quicker the result will be.

There are also concordancing programs that can run on the web. It may be more convenient to use web-based concordancing programs because they are accessible at any place and any time. These web-based concordancing programs are not affected as much by the speed of the individual learners' computers, as all the data processing is done on the server's side of the computer. Since the learners' computers may only receive text from the server, the bandwidth does not affect the display of results very much. The

factors that determine the time it takes to display results on the screen are the size of the corpora, the structure of the database, and the processing speed of the server.

There are two basic assumptions underlying a corpus-based approach to language teaching. First, corpus-based investigations show patterns and complexities that are normally overlooked in traditional intuition-based analyses. During this study, the subjects discovered subtle differences in meaning when words, especially verbs, were combined with a variety of collocates. This discovery then helped the subjects produce more accurate sentences. Secondly, students are encouraged to use contextual examples in authentic texts rather than referring to or depending on isolated examples in language textbooks. The subjects discovered that some words or phrases could be used with different collocates, depending on the context. Also depending on the context, some words may or may not be used. For example, the words “to watch” and “to monitor” have a similar meaning but are used within different contexts. The nature of language structure and use is demonstrated in corpus-based studies. Additionally, frequency of words and grammatical, structural and colloquial patterns are displayed in the data. The availability of such corpus-based analyses means that students’ learning needs, including patterns in construction and lexis, can be identified through the study of corpora. The compilation and study of corpora are made easier by text analysis software. Students are able to produce their own paradigms that fit the template that is specified by the search term. This results in a mixture of authentic text (newspapers, journals, magazines, etc.) and a concentration of information related to a paradigmatic format.

As the results of this study suggest, students can use a concordancing program as a tool to self-correct errors. Students may be able to use this tool instead of going to a writing center or meeting with a native speaker to proofread their writing or verify the grammar for at least sentence-level accuracy. Although they may get help at the writing center in terms of discourse level or organizational skills, a concordancing program can be a quick reference tool and is easily accessible. This may increase students’ confidence

level in their foreign language as they become independent from instructors and other resources.

Some of the subjects suggested that a concordancing program might have been more helpful if the corpora had consisted of the types of articles they read for their classes. This leads to another form of influence concordancing programs have on language teaching strategies, which includes evaluating and creating syllabi and materials. This can be especially helpful when creating an English for Specific Purposes course. Vocabulary, grammatical structures, and other information can be developed to meet the needs of the particular course. For example, a community college developing a course for certified nursing assistants for ESL learners needs to train the students not only in the subject matter, but also on language terms and phrases to be used in the classroom and on the job. Special online materials, such as those used in course management software, could also be made available to the students. In producing a course for specific subject matter aimed at a specific student population, materials developed with the help of corpora in concordancing programs can be based on the most up-to-date information available to the program's developers.

In addition to creating specialized materials based on corpora, creating subcorpora to be used in ESL classrooms, such as the one mentioned above, can be helpful in showing the semantic and syntactic contexts for words and phrases. Subcorpora are usually smaller in size and consist of homogeneous texts. An example of this implication in the study occurred when some of the subjects complained about the corpora (New York Times) being too difficult to understand, especially words that were used in political contexts. New words and expressions can be more manageable if corpora used in the concordancing program are used selectively. For example, subcorpora can be created to facilitate learning special vocabulary and grammatical structures that are used frequently or are specific to that area. ESL instructors can create learner corpora that can help define the most common errors ESL students make when they write in English.

In summary, concordancing programs support setting up environments where language learners have opportunities to interact with authentic texts to see how language is used in both general and particular settings. With a concordancing program, classes can be more learner-centered and inductive. While native-speaker intuition and introspective data in linguistics are the traditional resources used by native-speaking language teachers, corpus-based data has even more to offer: information on frequency, objectivity, a large quantity of data, ease of accessibility for nonnative speakers, details on style and register, absence of hesitation in giving a response or fatigue that could result from human informants, the availability of diachronic information, and the authenticity of the data.

5.6. Future research

This study has only investigated advanced ESL learners. In the future, it may be useful to know how lower- or intermediate-level foreign language learners utilize this program. While they probably will not recognize as many words as advanced learners, they may develop analytical strategies useful for finding sentence structures that are applicable to their research questions. In order to accommodate the learners' level of proficiency, future research can address the use of a parallel corpus. Because a parallel corpus shows examples in both the learners' L1 and the L2, learners may develop different strategies, depending on which language they are using to research. An advantage of using a parallel corpus is that learners can use search terms in either language, which may alter the way they use dictionaries and thesauri in conjunction with a concordancing program. By comparing examples in both languages that have the same meaning, learners can compare linguistic differences in those two languages. This helps them understand the context of the search word in the target language.

This study used only a general corpus. It would be helpful to know what the differences are in terms of the application of this program in writing when different

corpora are used. For example, the general belief is that the context of the corpus should be aligned with the goals and context of the writing. When using specialized corpora, writers may benefit more, as opposed to using corpora that are not specifically targeted to their area. Learners may be able to find more examples that match the tone and wording of the writing that is expected of them. As a result, they may be able to produce writing that is appropriate to the style of the subject matter. Learners may search not only for meanings of words, but also nuances. They may create different search patterns—search words, special characters, and the types of information they look for.

Lastly, this research was carried out over a short period of time, so the results do not show an improvement in writing in conjunction with subjects' grammatical knowledge before they were introduced to the concordancing program. This suggests that conducting a longitudinal study of foreign language learners may indicate the effects of using a concordancing program and the effects in terms of their progress in grammatical accuracy and users' strategies. A longitudinal study could also reveal students' attitudes toward using a concordancing program as they develop more strategies and are more familiar with the functions and special search characters. This information could prove to be invaluable, as this technology can change the face of language teaching and learning.

APPENDIX A
NEWSPAPER ARTICLE ON KOREAN PRESIDENT
IMPEACHMENT

Copyright 2004 *The Washington Post*

January 31, 2004 Saturday Final Edition

SECTION: A Section; A18

LENGTH: 550 words

HEADLINE: Roh Gets Taste of His Own Medicine; S. Korean Becomes Target of Anti-Graft Drive He Launched

BYLINE: Anthony Faiola and Joohee Cho, Washington Post Foreign Service

DATELINE: SEOUL

President Roh came to power promising to break the silence surrounding the estimated hundreds of millions of dollars of illegal funds funneled into election campaigns by South Korea's powerful family-owned conglomerates. During his 12 months in office, legal scholars say, Roh has granted unprecedented autonomy to prosecutors to launch the broadest corruption probes in national history.

Roh's efforts have led this new crop of reform-minded prosecutors straight to his doorstep.

Over the past six months, the prosecutors - now being hailed by corruption-weary South Koreans as national heroes, complete with their own online fan clubs - have forced the arrests of 16 politicians, aides and businessmen accused of accepting or handing out illegal cash to fund Roh's successful 2002 presidential bid. They include four top presidential aides, former campaign director and chauffeur.

Of the 273 lawmakers elected to the current National Assembly, 33 have been indicted on corruption charges, according to the prosecutor's office. "President ordered us to take no quarter, to be an independent force for justice, and that's just what prosecutors are doing,"

said one prosecutor involved in the investigations. He and other prosecutors interviewed declined to be named, citing the sensitive nature of the investigations and departmental policies. "We are finally free to do our jobs."

With the evidence mounting, Roh has conceded his campaign received illegal funds, but is now asking Koreans to forgive him. He argues that he is not as dirty as the opposition Grand National Party, charged with accepting tens of millions of dollars more in illegal cash than Roh's campaign during the 2002 elections. Roh has vowed to give up his job if his campaign is found to have been even fractionally as corrupt as the GNP, which has long maintained close ties to the conglomerates.

"If the amount of our illegal funds is more than one-tenth of Grand National Party's, I will resign from my presidential post and retire from politics," Roh said last month.

In the meantime, the National Assembly, controlled by the GNP, has appointed a special prosecutor to investigate illegal campaign financing at the Blue House, Korea's presidential palace. Korean law prohibits prosecutors from indicting a sitting president, but GNP leaders have threatened impeachment proceedings—an event that analysts say could precipitate a constitutional crisis at a time when neighboring North Korea is threatening to become the world's newest nuclear power.

The sweeping probes have shocked the nation in their scope, as politicians fled into hiding before turning themselves in and investigators raided Samsung, Hyundai Motors, SK Group, LG and the other corporate giants that have transformed Korea into an industrialized nation over the past 30 years. The raids have come after decades of tight control of the prosecutor's office by the executive branch.

Roh, prosecutors and legal experts say, can claim much of the credit for the newly independent prosecutors. Last March, in a televised debate with a group of state prosecutors, Roh challenged them to restore the nation's trust in them. By his side was Kang Kum Sil, a crusading former judge and human rights lawyer newly appointed by Roh to be South Korea's first female justice minister. Under her guidance, the prosecutor's office was reshuffled and aggressive, younger lawyers in their thirties and forties were assigned to key posts in the political corruption division.

APPENDIX B
NEWSPAPER ARTICLE ON NORTH KOREAN
NUCLEAR PROGRAM

Newspaper article on North Korean nuclear program

Copyright 2004 *The New York Times* Company

March 28, 2004, Sunday, Late Edition - Final

SECTION: Section 1; Page 11; Column 1; Foreign Desk

LENGTH: 550 words

HEADLINE: North Korea Rejects U.S. Demand to Scrap Its Nuclear Programs

BYLINE: By JOSEPH KAHN

DATELINE: BEIJING, March 27

North Korean radio on Saturday explicitly rejected the formula the United States has put forward as its bottom-line position in talks aimed at ending North Korea's nuclear programs, raising doubts about whether the fitful negotiations are making even limited progress.

The statement carried by Radio Pyongyang and monitored by news agencies in South Korea came just after a visit to North Korea by China's foreign minister, Li Zhaoxing, and shortly before a visit to the region by Vice President Dick Cheney that is planned for April. It used typically unrestrained language in accusing the United States of secretly planning a war.

"The present situation on the Korean peninsula remains dangerous owing to the reckless moves of the U.S. war hawks and their followers to unleash a war of aggression against the D.P.R.K. so that a nuclear war may break there anytime," it said, using the initials of North Korea's official name, the Democratic People's Republic of Korea.

Mr. Li said this week that North Korea was ready to "push forward" with a third round of talks involving the United States, South Korea, China, Japan and Russia. North Korea has

said it is willing to end its nuclear programs. But in the latest statement, it appears to be setting the stage for another inconclusive effort.

The statement rejected the American demand for a "complete, verifiable, irreversible dismantling" of the country's nuclear programs. Bush administration officials have repeatedly stated that they will not sign any agreement with North Korea that does not use that wording. The administration has also said it will not provide aid or other benefits to North Korea before it scraps all its nuclear programs and allows rigorous inspections.

While North Korea often harshly criticizes the United States for what it considers an inflexible stance, the Saturday announcement seemed to go further. It put North Korea on record as saying that it could not accept the main goals President Bush and his negotiators have insisted on in the first two rounds of talks.

The statement rejected the American formula point by point.

Complete nuclear dismantling is a plot to overthrow the North's socialist system after stripping it of its nuclear deterrent, it said.

"Verifiable nuclear dismantling reflects a U.S. intention to spy on our military capabilities before starting a war," it also said.

"Irreversible nuclear dismantling is nothing other than a noose to stifle us after eradicating our peaceful nuclear-energy industry," it said.

Bush administration officials have said that North Korean negotiators discussed accepting the American terms as a common negotiating goal in the last round of talks, held in

Beijing in February. North Korea did not do so in the end, however, and the talks concluded without even achieving China's goal of getting all the parties to accept a framework for future negotiations.

American officials say they believe that North Korea has already produced at least two nuclear bombs and could make many more.

North Korea has demanded that the United States make concessions, including providing energy aid and pledging not to use force against North Korea, before it fully dispenses with what it calls its nuclear deterrent.

The two sides also disagree about the American demand that an accord eliminate all nuclear programs, including those North Korea contends are intended to produce nuclear power for civilian uses.

APPENDIX C
PLANS FOR PARAPHRASING SESSIONS

You will paraphrase a newspaper article. You are allowed to use *MonoConc Pro*, *Yahoo* online dictionaries and an online thesaurus. Use Microsoft Word for writing. Your screen will be recorded during the writing session. After you have finished writing, you and the researcher will watch the video together for stimulated recall.

Please ask me now if you have any questions.

APPENDIX D
DIRECTIONS FOR PARAPHRASING

Paraphrase the newspaper article. Your writing should be double-spaced and half a page or longer. Your writing should include important information from the article. You are allowed to use *MonoConc Pro* and *Yahoo* online dictionaries and thesaurus.

There is no time limit.

When you paraphrase, you may restate the ideas of the newspaper. Use your own grammar structure and words.

Avoid using the same sentence structure or only changing some of the words. Do not change the meaning of the newspaper article.

Please ask me now if you have any questions.

APPENDIX E
DIRECTIONS FOR STIMULATED RECALL SESSION

Directions for Stimulated Recall Session

We are going to watch the video now. I'm interested in what you were thinking at the time you were paraphrasing the article. We can see what you were searching for in *MonoConc Pro*, online dictionaries, and thesaurus. We can also see what you were writing, but I don't know what you were thinking. So, I'd like you to tell me what you were thinking, what was on your mind, and what you are reading at that time while you were using *MonoConc*, the *Yahoo* dictionaries and thesaurus.

You can tell me to pause the video at any time. If you want to tell me something about what you were thinking, tell me to pause the video. If I have a question about what you were thinking, then I will pause the video and ask you to talk about that part of the video.

Please ask me now if you have any questions.

APPENDIX F
DEMOGRAPHIC DATA

This questionnaire is taken and modified from Yoon and Hirvela (2004).

1. Name _____
2. Email address _____
3. Gender: ___Male ___Female
4. Graduate Program _____
5. Major _____
6. How long have you been in the US or other English speaking countries?
_____ years _____ months
7. How long have you been in colleges in the US? _____ years
_____ months
8. In general, do you like to use a computer? ___Yes ___No
9. How often do you use a computer for personal purposes (e.g., email, chat)?
(Check one)
___Several times or more a day ___About once a day ___Several times a week
___About once a week ___About once a month ___Seldom
___Other (list)
10. How often do you use a computer for schoolwork (e.g., writing a paper)? (Check one)
___Several times or more a day ___About once a day
___Several times a week ___About once a week ___About once a month
___Seldom ___Other (list)
11. When you use a computer for personal purposes, do you use English or your native language?
___English ___Native Language ___Both ___Other (list)

12. How much do you use Korean when using a computer?

☐ Almost none ☐ About 25% ☐ About 50%

☐ About 75% ☐ Others (list)

13. Do you have Internet access at home? ☐ Yes ☐ No

14. Do you use any dictionaries for writing in English? ☐ Yes ☐ No

15. If yes, what kind of dictionaries do you often use? Check all that apply.

☐ English-Korean ☐ Korean-English ☐ English-English

☐ Paper dictionary ☐ On-line dictionary ☐ Dictionary software

16. Do you use a thesaurus for writing in English? Check all that apply.

☐ Paper thesaurus ☐ On-line thesaurus ☐ Thesaurus software

17. Have you used the Internet as a reference for your composing and/or revising

papers or any types of writing? ☐ Yes ☐ No

18. If yes, explain what and how you used it.

19. Had you heard about corpora before you participated in this study?

☐ Yes ☐ No

20. If yes, had you used corpora before you participated in this study?

☐ Yes ☐ No

21. If yes, which corpus did you use? List all _____

APPENDIX G
SURVEY

This survey is taken and modified from Yoon and Hirvela (2004).

Reaction to using the *English Gigaword* (NYT from year 2000-2002 only) corpus

The following questions are regarding your opinions on using the corpus provided on the CD ROM. Please use the scale below to circle the response that most closely resembles your perspectives.

1 (strongly agree) ----- 6 (strongly disagree) 0=No Opinion

1. The searching technique was easy to learn.	1 2 3 4 5 6 N
2. The practice session last week was helpful for learning the techniques.	1 2 3 4 5 6 N
3. The corpus is more helpful than a dictionary for my English writing.	1 2 3 4 5 6 N
4. Using the corpus is helpful for finding out the meaning of vocabulary.	1 2 3 4 5 6 N
5. Using the corpus is helpful for finding out the usage of vocabulary.	1 2 3 4 5 6 N
6. Using the corpus is helpful for finding out the usage of phrases.	1 2 3 4 5 6 N
7. Using the corpus is helpful for finding out grammar.	1 2 3 4 5 6 N
8. Using the corpus improved my English writing skills.	1 2 3 4 5 6 N
9. Using the corpus increased the grammatical accuracy of my writing in English.	1 2 3 4 5 6 N
10. I have some difficulty in using the corpus due to limited access to computers.	1 2 3 4 5 6 N
11. I have some difficulty in using the corpus due to time and effort spent on analyzing the data.	1 2 3 4 5 6 N
12. I have some difficulty in using the corpus due to unfamiliar	1 2 3 4 5 6 N

vocabulary in concordancing output.

13. I have some difficulty in using the corpus due to cut-off sentences in concordancing output.	1 2 3 4 5 6 N
14. I have some difficulty in using the corpus due to too many sentences in concordancing output.	1 2 3 4 5 6 N
15. I have some difficulty in using the corpus due to the limited number of sentences in concordancing output.	1 2 3 4 5 6 N
16. I have some difficulty in analyzing concordancing output.	1 2 3 4 5 6 N
17. I have some difficulty in analyzing collocation frequencies.	1 2 3 4 5 6 N
18. I have some difficulty in performing the search techniques.	1 2 3 4 5 6 N
19. The real texts in the corpus are too difficult to understand.	1 2 3 4 5 6 N
20. I understand the purpose of using the corpus for my writing.	1 2 3 4 5 6 N
21. I want to use the corpus for writing in English again.	1 2 3 4 5 6 N
22. When I have problems writing in English, I search for help in the corpus.	1 2 3 4 5 6 N
23. When I search for information in the corpus, I usually get the information that I need.	1 2 3 4 5 6 N
24. I use the corpus when writing for purposes other than academic writing.	1 2 3 4 5 6 N
25. As I have learned more about the corpus, I have come to like it more.	1 2 3 4 5 6 N
26. I will use the corpus for writing in English in the future.	1 2 3 4 5 6 N
27. Learning about the corpus has increased my confidence about writing in English.	1 2 3 4 5 6 N
28. If I had used the corpus earlier, I would have had more accurate papers or other types of writing.	1 2 3 4 5 6 N
29. Overall, the corpus is a very useful resource for writing in	1 2 3 4 5 6 N

English.	
30. Corpus use is more helpful for writing than for reading in English.	1 2 3 4 5 6 N
31. The corpus should be introduced to other international students and ESL students.	1 2 3 4 5 6 N
32. I recommend using the corpus in the ESL writing courses in the future.	1 2 3 4 5 6 N
33. Corpus use should be taught in English classes in Korea.	1 2 3 4 5 6 N
34. I will recommend the corpus to other international students at the University of Iowa or elsewhere.	1 2 3 4 5 6 N

The following questions are about your experience using the corpus last week.

1. Last week, I used the corpus (check one)

___ never ___ once ___ several times ___ once every day ___ several times every day

2. I used the corpus for (check all that apply)

___ writing assignments ___ emailing ___ memos ___ others (explain)

3. I used the corpus (check all that applies)

___ at school ___ at home ___ other places (explain)

4. I used the corpus to look for the following kinds of words (check all that apply);

___ nouns ___ verbs ___ adjectives ___ adverbs

___ prepositions ___ phrases ___ others (explain)

5. I used the corpus to see the following information (check all that apply);

___ sample sentences ___ context ___ collocations ___ word frequencies ___ collocation frequencies ___ others (explain)

6. I used the following special characters (check all that apply);

___ * ___ @ ___ ? ___ others (explain)

APPENDIX H
IRB DOCUMENTS: INFORMED CONSENT
DOCUMENTS

Project Title: Using a corpus as a reference tool for English for academic purposes

Researcher: Kyosung Koo

WHAT IS THE PURPOSE OF THIS STUDY?

This is a research study. We are inviting you to participate in this research study because you are a high-advanced nonnative speaker of English from Korea enrolled in graduate programs at the University of Iowa.

The purpose of this research study is to find out how English as a second language speakers from Korea use reference tools such as dictionaries when paraphrasing a newspaper article.

HOW MANY PEOPLE WILL PARTICIPATE?

Approximately 10 people will take part in this study at the University of Iowa.

HOW LONG WILL I BE IN THIS STUDY?

If you agree to take part in this study, your involvement will last for about 9 days. We will meet 2 times, with an interval of a week between the two meetings. For the first meeting you will be required to participate approximately 1 hour. We will then have the second meeting a week later. You will be required to participate approximately 1.5 hours (1 hour and 30 minutes).

WHAT WILL HAPPEN DURING THIS STUDY?

There will be two meetings, and these meetings will take approximately 2.5 hours total. There will be a week gap between the two meetings.

1. At the first meeting, you will be asked to answer general questions regarding using online and paper-version dictionaries, computers, the Internet, and a corpus.
2. You will be learning basic text-searching functions of a concordancing program called *MonoConc* including loading a corpus, frequency of occurrences, single- and multiple-word searches, excluding words from consideration, using wildcards, and displaying contexts.
3. You will be asked to do the following activities using a concordancing program on a PC: Finding out the word that occurs the most frequently in the corpus, searching and displaying a word and its surrounding words, searching and displaying multiple words and their surroundings, identifying prepositions used with a certain verb, and figuring out different meanings of synonyms.
4. You will then be asked to use this program for your own writing purposes for a week in order to get used to using this program.
5. A week later, there will be another meeting for paraphrasing a newspaper article.
6. You will be given one newspaper article. You can use online dictionaries and a concordancing program during the tasks. The newspaper article will be 2-3 paragraphs long on a current issue that you may be familiar with. There will not be a time limit for paraphrasing the article.

WHAT ARE THE RISKS OF THIS STUDY?

There are no foreseeable risks to participating.

WHAT ARE THE BENEFITS OF THIS STUDY?

You will not benefit from being in this study. However, we hope that, in the future, other people might benefit from this study because ESL teachers can adapt the results to create more effective and meaningful learning conditions. ESL students will then be able to use authentic texts that are generally the most useful for learners. By being exposed to

authentic texts, students can expand their understanding of word functions in particular contexts.

WILL IT COST ME ANYTHING TO BE IN THIS STUDY?

You will not have any costs for being in this research study.

WILL I BE PAID FOR PARTICIPATING?

You will be paid for being in this research study. You will be paid an average of \$15 for each meeting. When you complete the second meeting, you will be paid a total of \$30.

WHO IS FUNDING THIS STUDY?

The University and the research team are receiving no payments from other agencies, organizations, or companies to conduct this research study.

WHAT ABOUT CONFIDENTIALITY?

We will keep your participation in this research study confidential to the extent permitted by law. However, it is possible that other people may become aware of your participation in this study. For example, federal government regulatory agencies and the University of Iowa Institutional Review Board (a committee that reviews and approves research studies) may inspect and copy records pertaining to this research. Some of these records could contain information that personally identifies you.

To help protect your confidentiality, I will assign a nickname in order to remain anonymous. A master list of names and nicknames will be stored in a text format on the researcher's computer hard drive. The text file will be password protected. The data and the master list will be retained until the transcription is finished. All confidential data will be deleted when data transcription is done.

If we write a report or article about this study or share the study data set with others, we will do so in such a way that you cannot be directly identified.

Audio/Video Recording or Photographs

One aspect of this study involves making audio recordings of you. These recordings will be necessary to gain your thought process during writing and to analyze using reference tools. I, the researcher, will only have access to the recordings. All the recordings will be destroyed as soon as the data analysis is done.

☐ Yes ☐ No I give you permission to make [audio recordings / video recordings / photographs] of me during this study.

IS BEING IN THIS STUDY VOLUNTARY?

Taking part in this research study is completely voluntary. You may choose not to take part at all. If you decide to be in this study, you may stop participating at any time. If you decide not to be in this study, or if you stop participating at any time, you won't be penalized or lose any benefits for which you otherwise qualify.

WHAT IF I HAVE QUESTIONS?

We encourage you to ask questions. If you have any questions about the research study itself, please contact: Kyosung Koo (researcher)/(319) 335-6465 or James Pusack (supervising faculty) at (319) 335-2203.

If you have questions about the rights of research subjects or research related injury, please contact the Human Subjects Office, 300 College of Medicine Administration Building, The University of Iowa, Iowa City, Iowa, 52242, (319) 335-6564, or e-mail

irb@uiowa.edu. General information about being a research subject can be found by clicking “Info for Public” on the Human Subjects Office web site, <http://research.uiowa.edu/hso>.

Your signature indicates that this research study has been explained to you, that your questions have been answered, and that you agree to take part in this study. You will receive a copy of this form.

Subject's Name (printed): _____

(Signature of Subject)

(Date)

REFERENCES

- Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. H. Granger, J. Petch-Tyson, S. (Ed.), *Computer learner corpora, second language acquisition and foreign language teaching* (Vol. 6, pp. 55-76). Philadelphia: John Benjamins.
- Anderson, R. C., & Shiffrin, Z. (1980). The meaning of words in context. In R. Spiro, B. Bruce & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 331-348). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ard, J. (1982). The use of bilingual dictionaries by ESL students while writing. *ITL Review of Applied Linguistics*, 58, 1-27.
- Aston, G. (1996). The British national corpus as a language learner resource. In S. Botley, J. Glass, T. McEnery & A. Watson (Eds.), *Proceedings of teaching and language corpora* (pp. 178-191). Lancaster, UK: University Centre for Computer Corpus Research on Language.
- Aston, G. (1997a). Enriching the learning environment: Corpora in ELT. In S. F. A. Wichmann, T. McEnery, G. Knowles (Ed.), *Teaching and language corpora* (pp. 51-64). New York: Longman.
- Aston, G. (1997b). Small and large corpora in language learning. In B. Lewandowska-Tomaszyk & P. J. Melia (Eds.), *Practical applications in language corpora* (pp. 51-62). Lodz, Poland: University Press.
- Barlow, M. (1996b). Parallel texts in language teaching. In J. G. S. Botley, T. McEnery, A. Wilson (Ed.), *Proceedings of teaching and language corpora* (Vol. 9, pp. 45-56). Lancaster: UCREL Technical Papers.
- Barlow, M. (2001). *ParaConc*. Houston, TX: Athelstan.
- Barlow, M. (2002). *MonoConc Pro* (Version 2.2). Houston: Athelstn.
- Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? *TESOL Quarterly*, 14, 325-336.
- Bejoint, H. (1981). The foreign student's use of monolingual English dictionaries. *Applied Linguistics*, 2, 207-222.
- Bishop, G. (2000). Dictionaries, examinations and stress. *Language Learning Journal*, Summer 2000(21), 57-65.
- Carrell, P. L., Devine, J., & Eskey, D. E. (1988). *Interactive approaches to second language reading*. Cambridge; New York: Cambridge University Press.
- Carter, R. (1993). Language awareness and language learning. In M. Hoey (Ed.), *Data description, discourse* (pp. 139-150). London: Harper-Collins.
- Carter, R., & McCarthy, M. (Eds.). (1988). *Vocabulary and language teaching*. New York, NY: Longman.

- Chandler, B. (1989). *Longman mini-concordancer*. London: Longman.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301-315.
- Curado, A. (2002). Exploitation and assessment of a business English corpus through language learning tasks. *ICAME Journal: Computers in English Linguistics*, 26, 5-32.
- Danielsson, P., & Ridings, D. (1996). Corpus and terminology: Software for the translation program at Göteborgs Universitet or getting students to do the work. In *Proceedings of teaching and language corpora 1996 (talc'96)* (pp. 57-67). Lancaster: UCREL.
- DiPardo, A. (1994). Stimulated recall in research on writing: An antidote to "I don't know, it was fine". In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (Vol. 8, pp. 163-181). Thousand Oaks, CA: Sage.
- Dodd, B. (1997). Exploiting a corpus of written German for advanced language learning. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora*. New York, NY: Longman.
- Dubin, F., Eskey, D. E., & Grabe, W. (Eds.). (1986). *Teaching second language reading for academic purposes*. Reading, MA: Addison-Wesley.
- Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence from 11 readers. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181-202). Norwood, NJ: Ablex.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21(2), 231-244.
- Flowerdew, J. (1996). Concordancing in language learning. In M. Pennington (Ed.), *The power of call* (pp. 97-113). Houston, TX: Athelstan.
- Francis, G. (1993). A corpus driven approach to grammar—principles, methods and examples. In M. Baker, G. Francis & E. Tognini-Bonelli (Eds.), *Text and technology* (pp. 137-156). Amsterdam/Philadelphia: Benjamins.
- Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *The Modern Language Journal*, 87(2), 168-199.
- Freser, C. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21, 225-241.
- Fuentes, A. C., & Rokowski, P. E. (2002). *Use of corpus in the business English classroom: From lesser to more specific* (Paper presented at the IVACS Conference). Limerick, Ireland: (ERIC Document Reproduction Service No. ED 477 551).
- Garner, R. (1988). Verbal-report data on cognitive and metacognitive strategies. In C. E. Weinstein, E. T. Goetz & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction and evaluation* (pp. 63-76). New York, NY: Academic Press.
- Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Gavioli, L. (1997). Exploring texts through the concordancer: Guiding the learner. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 83-99). London, UK: Longman.
- Gipe, J. P., & Arnold, R. D. (1979). Teaching vocabulary through familiar associations and contexts. *Journal of Reading Behavior*, 11(3), 282-285.
- Gonzalez, O. (1999). Building vocabulary: Dictionary consultation and the ESL student. *Journal of Adolescent & Adult Literacy*, 43(3), 264-270.
- Graff, D. (2003). *English Gigaword* [text]. Philadelphia, PA: Linguistic Data Consortium.
- Harvey, K., & Yuill, D. (1997). A study of the use of a monolingual pedagogical dictionary by learners of English engaged in writing. *Applied Linguistics*, 18(3), 253-278.
- Haynes, M. (1993). Patterns and perils of guessing in second language reading. In M. H. T. Huckin, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 46-64). Norwood, NJ: Ablex.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud & H. B. Joint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan Academic & Professional Ltd.
- Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80, 327-339.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Jabbour, G. (2001). *Corpus based instruction*: (ERIC Document Reproduction Service No. ED 452 734).
- James, C. (1983). Cloze procedures and long-term vocabulary retention. In K. M. Bailey, M. H. Long & S. Peck (Eds.), *Second language acquisition studies* (pp. 199-213). Rowley, MA: Newbury House.
- Johns, T. (1988). Whence and whither classroom concordancing? In T. Bongaerts, P. d. Haan, S. Lobbe & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9-33). Dordrecht: Foris Publications.
- Johns, T. (1991a). Should you be persuaded: Two examples of data driven. *ELR Journal*, 4, 1-16.
- Johns, T. (1991b). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *English Language Research Journal*, 4, 27-45.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Approaches to pedagogic grammar* (pp. 293-313). Cambridge, UK: Cambridge University Press.

- Johns, T. (2002). *Data-driven learning: The perpetual challenge*. Paper presented at the Teaching and learning by doing corpus analysis, Graz 19-24 July, 2000.
- Kelly, P. (1990). Guessing: No substitute for systematic learning of lexis. *System*, 18, 199-207.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, 5(3), 77-90.
- Kennedy, G. (1987a). Expressing temporal frequency in academic English. *TESOL Quarterly*, 21, 69-86.
- Kennedy, G. (1987b). Quantification and the use of English: A case study of one aspect of the learner's task. *Applied Linguistics*, 8, 264-286.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. New York, NY: Longman.
- Knight, S. (1994). Dictionary use while reading: The effect on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, 78(3), 285-299.
- Laufer, B. (1997). The lexical plight in second language reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20-34). Cambridge: Cambridge University Press.
- Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and "bilingualised" dictionaries in the comprehension and production of new words. *The Modern Language Journal*, 81, 189-196.
- Laufer, B., & Kimmel, M. (1997). Bilingualised dictionaries: How learners really use them. *System*, 25, 361-369.
- Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, 28(1), 89-108.
- Laufer, B., & Sim, D. D. (1985). Taking the easy way out: Non-use and misuse clues in efl reading. *English Teaching Forum*, 23(2), 7-10, 20.
- Leech, G. (1997). Teaching and language corpora - a convergence. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). New York, NY: Longman.
- Lewis, M. (1993). *The lexical approach: The state of ELT and the way forward*. Hove, UK: Language Teaching Publication.
- Lewis, M. (1997). Pedagogical implication of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 255-270). Cambridge: Cambridge University Press.
- Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16, 33-42.

- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29(6), 861-878.
- Ma, K. C. (1993). *Classroom concordancing: How and what students learn*. Unpublished M.A. dissertation, City Polytechnic of Hong Kong, Hong Kong.
- Mackintosh, K. (1998). An empirical study of dictionary use in L2-L1 translation. In B. T. S. Atkins (Ed.), *Using dictionaries: Studies of dictionary use by language learners and translators* (pp. 123-149). Tübingen: Max Niemeyer Verlag.
- Maddalena, S. R. (2001). An investigation into how corpus analysis may be used in the second language classroom to solve some of the problems surrounding non-native speakers' understanding of seemingly synonymous words. (ERIC Document Reproduction Service No. ED 458795).
- McDonough, S. H. (1986). *Psychology in foreign language teaching* (2nd ed.). London: Routledge.
- McEnery, T., & Wilson, A. (1993). The role of corpora in computer-assisted language learning. *Computer Assisted Language Learning*, 6(3), 233-248.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics: An introduction* (2nd ed.). Edinburgh, UK: Edinburgh University Press.
- McEnery, T., Wilson, A., & Baker, J. (1997). Teaching grammar again after twenty years: Corpus-based help for teaching grammar. *ReCALL*, 9(2), 8-16.
- Mintz, T., Newport, E., & Bever, T. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, 26(4), 393-424.
- Minugh, D. (1997). All the language that's fit to print: Using British and American newspaper CD-ROMs as corpora. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 67-82). New York, NY: Longman.
- Moudraia, O. (2001). *Lexical approach to second language teaching*. Washington DC. (ERIC Document Reproduction Service No. ED 455698).
- Murphy, B. (1996). Computer, corpora and vocabulary study. *Language Learning Journal*, 14, 53-57.
- Möllering, M. (2001). Teaching German modal particles: A corpus-based approach. *Language Learning & Technology*, 5(3), 130-151.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge; New York: Cambridge University Press.
- Nattinger, J. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, 14, 337-344.
- Nist, S. L., & Olejnik, S. (1995). The role of context and dictionary definitions on varying levels of word knowledge. *Reading Research Quarterly*, 30, 172-193.

- Paikeday, T. M. (1993). The Kenkyusha dictionaries: 50 years of shortchanging the Japanese. *JALT Journal*, 15(2), 239-244.
- Paribakht, T. S., & Wesche, M. (1999). Reading and incidental L2 vocabulary acquisition. *Studies in Second Language Acquisition*, 21, 195-224.
- Parry, K. (1991). Building a vocabulary through academic reading. *TESOL Quarterly*, 25, 629-651.
- Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109-129). Norwood, NJ: Ablex.
- Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, 80, 478-493.
- Rutherford, W. E. (1987). *Second language grammar: Learning and teaching*. London, UK: Longman.
- Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, 54(3), 525-582.
- Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, 21, 439-453.
- Scott, M. (2000a). Reverberations of an echo. In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *PALC'99: Practical applications in language corpora. Lodz studies in language* (Vol. 1, pp. 49-68). Frankfurt: Peter Lang.
- Scott, M. (2000b). *Wordsmith tools* (Version 3.0). Oxford, UK: Oxford University Press.
- Scott, M., & Johns, T. (1993). *Microconcord* (Version 1.0). Oxford: Oxford University Press.
- Sinclair, J. (1986). Basic computer processing of long texts. In G. Leech & C. N. Candlin (Eds.), *Computers in English language teaching and research* (pp. 185-203). Harlow, Essex: Longman.
- Sinclair, J. (1987). *Looking up: An account of the COBUILD Project in lexical computing*. London, UK: Collins COBUILD.
- Sinclair, J. (Ed.). (1990). *Collins COBUILD English grammar*. London: Harper Collins.
- Smagorinsky, P. (1994). Think-aloud protocol analysis: Beyond the black box. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (Vol. 8, pp. 3-19). Thousand Oaks, CA: Sage.
- St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning & Technology*, 5(3), 185-203.
- Stein, M. (1993). The healthy inadequacy of contextual definition. In M. H. T. Huckin, & J. Coady (Ed.), *Second language reading and vocabulary learning* (pp. 203-212). Norwood, NJ: Ablex.

- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stevens, V. (1990). Text manipulation: What's wrong with it, anyway? *CAELL Journal*, 1(2), 5-8.
- Stevens, V. (1991b). Concordance-based vocabulary exercises: A viable alternative to gap-filling. *ELR Journal*, 4, 47-61.
- Stevens, V. (1995). Concordancing with language learners: Why? When? What? *CAELL Journal*, 6(2), 2-10.
- Thompson, P. (2002). What use are corpora in the teaching of EAP. In G. Aston (Ed.), *Proceedings of the fifth teaching and language corpora conference*. Bertinoro, Italy: University of Bologna.
- Thurston, J., & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes Journal*, 17(3), 267-280.
- Todd, R. W. (2001). Induction from self-selected concordances and self-correction. *System*, 29(1), 91-102.
- Tribble, C. (1990). Concordancing and an EAP writing program. *CAELL Journal*, 1(2), 10-15.
- Tribble, C. (2001). Small corpora and teaching writing: Towards a corpus-informed pedagogy of writing. In M. Ghadessy, A. Henry & R. Roseberry (Eds.), *Small corpus studies and ELT: Theory and practice* (pp. 381-408). Philadelphia, PA: John Benjamins.
- Tribble, C., & Johns, G. (1990). *Concordances in the classroom*. Harlow, UK: Longman.
- Turnbull, J., & Burston, J. (1998). Towards independent concordance work for students: Lessons from a case study. *ON-CALL*, 12(2), 10-21.
- Varantola, K. (1998). Translators and their use of dictionaries. In B. T. S. Atkins (Ed.), *Using dictionaries: Studies of dictionary use by language learners and translators* (pp. 179-192). Niemeyer: Tübingen.
- Wang, L. (2000). English-Chinese parallel concordancer. Birmingham, UK: University of Birmingham.
- Wang, L. (2001). Exploring parallel concordancing in English and Chinese. *Language Learning & Technology*, 5(3), 174-184.
- Wichmann, A. (1995). Using concordances for the teaching of modern languages in higher education. *Language Learning Journal*, 11, 61-63.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London, UK: Collins COBUILD.

- Yoon, H. (2005). *An investigation of students' experiences with corpus technology in second language academic writing*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257.
- Zanettin, F. (1994). Parallel words: Designing a bilingual database for translation activities. In A. Wilson & T. McEnery (Eds.), *Corpora in language education and research: A selection of papers from talc94* (Vol. 4, pp. 99-111). Lancaster, UK: UCRE.