AN INVESTIGATION OF STUDENTS’ EXPERIENCES WITH CORPUS TECHNOLOGY IN SECOND LANGUAGE ACADEMIC WRITING

DISSertation

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Hyunsook Yoon, M.A.

* * * * *

The Ohio State University
2005

Dissertation Committee:  Approved by

Dr. Shelley D. Wong, Adviser  ______________________

Dr. Alan R. Hirvela, Adviser  Adviser

Dr. Charles R. Hancock  ______________________

Dr. Jack H. Rouzer  Adviser

College of Education
Copyright by
Hyunsook Yoon
2005
ABSTRACT

An increasing number of recent studies have applied the insights and methodologies of corpus linguistics to second language (L2) writing pedagogy and research. These studies have generated a highly theoretical understanding of corpus use in L2 writing, but still leave gaps about the practical application for L2 writing teachers. This study examined the quality of students’ writing experiences with corpus use, how it affects their development of L2 competence, and most importantly, how corpus technology can be integrated into L2 writing instruction. The research site of the study was a graduate level ESL academic writing course at a large midwestern American university. Six students participated in an in-depth case study to provide rich descriptions of corpus pedagogy in L2 writing. This qualitative research involved a wide array of ethnographic techniques, including: classroom observations, open-ended interviews, multiple case studies, and a grounded survey.

The findings revealed that corpora served as a meaningful reference for language input by providing students with common usage and collocation patterns. More importantly, corpus use seemed to have an immediate effect by helping them to solve immediate writing/language problems, and also to develop potentially longer-term, cognitive skills by promoting their perceptions of lexico-grammar and language awareness. A significant finding is that the corpus experience seemed to promote
independent learning (i.e., students took more responsibility for their own writing). As the corpus approach was introduced and linked to the writing process, the students became more independent writers through increased access to linguistic resources. By doing so, they approached L2 writing with more ease, and their overall confidence in writing increased.

This study identified a wide variety of individual experiences and learning contexts that were involved in deciding the levels of the students’ willingness and success in using corpora. The findings suggest that writing teachers who attempt to incorporate the corpus linguistics approach into their instruction need to understand the multi-faceted aspects of technology use that could facilitate or impede the individual student’s L2 writing.
Dedicated to my parents
ACKNOWLEDGMENTS

A dissertation is not a lone enterprise of a solitary scholar, but it requires the help, encouragement and support of many people. First and foremost, I would like to thank all the students who were my supreme teachers in this project. We shared many hours in and out of the classroom exploring the many facets of learning second language writing in its native context. They generously – and at times courageously – opened their personal lives to me to share their experiences and insights. All of us were rewarded with the joy of discovery. In the end, I can call them my valued colleagues, but most of all, they are my good friends. In some small way, I hope this dissertation contributes to their hopes: to pave the journey for future learners to become more proficient L2 writers.

I am deeply grateful to the members of my dissertation committee who guided me through the rigors of conducting research for the completion of my doctoral program. My first thanks goes to my academic adviser, Dr. Shelley Wong, who has provided me with intellectual inspiration and emotional support through my years of graduate studies at OSU. Her passion about teaching is indelibly cherished in my heart. She gave me compassionate guidance, thoughtful insight, and support in countless ways too many to mention here. My special thanks goes to my co-adviser, Dr. Alan Hirvela, whose intellectual mentorship enabled me to advance my thinking about approaches to L2 writing during my graduate career. His caring support and intellectual passion have been
beacons of enlightenment along the way, contributing immensely to my professional development and future career. I also wish to thank Dr. Charles Hancock for his precious time and careful reading of my written product. This study benefited a lot from his critical judgment that sharpened my analytical and critical thinking skills. I also owe a deep debt of gratitude to Dr. Jack Rouzer who shared his classroom, students, and also opened an exciting door to the world of corpus pedagogy in L2 writing. This study owes a great deal to his initiative, hospitality, and expertise.

I am also grateful to Dr. Diane Belcher, now at Georgia State University, for her moral support and intellectual motivation. She counseled me with words of wisdom when I was struggling to find a direction for this study. I would like to extend special thanks to Dr. Evelyn Freeman, now Dean of the OSU campus at Mansfield, for her emotional and financial support so instrumental for my studies at OSU. Also, she showed me how one person can make a big difference in the world. I owe a great deal to Ms. Lisa Margeson, who was my supervisor in my graduate associate position for a couple of years. She showed patience, understanding and flexibility in helping me balance my work with my studies. Also, she was a cultural resource who explained to me the perplexities of life in this country.

There are also many colleagues and friends who walked the long road of graduate school with me, giving me the needed support and encouragement when travel got weary. Above everyone else, I would like to acknowledge Dr. Stephen Zolovinski for his tireless encouragement as well as his proofreading that was instrumental in getting this product completed. His support continued after he graduated from the OSU Department of
Anthropology to take on a post-doctoral position in the Philippines. His phone calls from OSU and abroad always kept me on track in getting the final draft written.

I have been very fortunate to have wonderful colleagues, Hyungmi Joo and Youngjoo Yi. We shared the highs of excitement and lows of frustration as we worked to complete our studies. Their companionship is priceless, and it will be cherished throughout the rest of my life. My heartfelt gratitude goes to a former OSU graduate student, Dr. En-Chong Liaw, now at a Taiwan university for her friendship. We shared long hours together at the library during the hardest time of my writing, boosting me with support and energy. I also wish to thank Yuka Kurihara for her intellectual discussion, invaluable encouragement and ever-present smiles. She taught me to see things differently. And many thanks to Dr. Seung Jung Kim, Mihye Seo, Jiwon Seo, Moon-kyung Kim, and Heejei Yoon for their friendship and caring, which made my life in Columbus a pleasant memory. I thank them all.

I would like to express great appreciation to my family. This dissertation is dedicated to my parents, Won-ik Yoon and Jung-soon Song, who have always had the faith, trust, and hope in me to make my dreams come true. To them I give a big round of applause. Also, special thanks to my brother, Dae-hyun, and sisters, Mi-ye, Soon-sik and Kyung-sook, whose support kept me going as I pursued my studies in a foreign country.

Finally, but not least, I am most grateful to my husband, Dr. Bang-Chool Kim, for his continued support and inspiration both when he was an OSU graduate student and now as a scholar in Korea. As my best friend, mentor, and constant source of encouragement, he often believed in me more than I did myself.

To all the above, I owe an invaluable debt.
VITA

July 15, 1970………………………………………Born – Chongju, Korea

1994………………………………………………..B.A. English,  
Hankuk University of Foreign Studies

1996………………………………………………..M.A. Communications,  
Korea University

1999-2000…………………………………………Assistant Teacher,  
Korea National Open University

2001………………………………………………..Vice President,  
Council of Graduate Students  
The Ohio State University

2001-present……………………………………..Graduate Associate,  
The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Education  
Foreign and Second Language Education; TESOL
Minor Fields: Second Language Literacy; Research Methodology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
</tbody>
</table>

Chapters:

1. Introduction
   1.1 Background of the study.
   1.2 Statement of the problem.
   1.3 Purpose of the study.
   1.4 Research questions.
   1.5 Overall research methodology.
   1.6 Significance of the study.
   1.7 Definitions of key terms.
   1.8 Assumptions of the study.

2. Review of the Literature
   2.1 Corpus linguistics and its view of language.
   2.2 Collocation and L2 (writing) proficiency.
   2.3 Controversies about using the corpus approach in language teaching.
   2.4 Corpus studies in the development of classroom materials and activities.
   2.5 Research on students’ use of corpus technology in L2 writing.

3. Methodology
   3.1 Initial field experience.
   3.2 Design of the study.
   3.3 Context of the study.
       3.3.1 Site.
       3.3.2 Corpus.
   3.4 Participants.
3.4.1 Student writers………………………………………………………60
3.4.2 Instructor…………………………………………………………….62
3.5 Data collection…………………………………………………………..65
3.5.1 Participant observation………………………………………………65
3.5.2 Case study……………………………………………………………67
3.5.3 Survey…………………………………………………………………70
3.6 Data analysis……………………………………………………………..71
3.7 Trustworthiness and ethics of the study………………………………73

4. Six Cases Writers and Language Learners: Portraits…………………………77
4.1 Introducing case study students……………………………………………….77
4.2 Views on language and language learning…………………………………99
4.3 Student approaches to L2 writing………………………………………….106
4.4 Difficulties in L2 writing…………………………………………………..114
4.5 L2 writing processes………………………………………………………..120
4.6 Learning styles……………………………………………………………..129
4.7 Summary………………………………………………………………….134

5. Student Use of Corpus Technology in L2 Writing……………………………..137
5.1 Students’ corpus use in L2 writing……………………………………………138
  5.1.1 Corpus use patterns and benefits………………………………………..138
  5.1.2 Corpus use, problems/difficulties………………………………………..175
  5.1.3 Students’ overall evaluation of corpus use in L2 writing……………….186
5.2 Influence of corpus use on their understanding of language and L2 writing.191
  5.2.1 Changes in their understanding of language and language learning..192
  5.2.2 Changes in their approaches to L2 writing and the writing process..204
5.3 Students’ perspectives on collocation in L2 learning and writing………..222
  5.3.1 Role of L1 in L2 collocation uses……………………………………….222
  5.3.2 Collocation and L2 writing…………………………………………….226

6. Discussion and Conclusion…………………………………………………………233
6.1 The role of corpus use in L2 academic writing………………………………233
  6.1.1 Students’ corpus use in L2 writing………………………………………..234
  6.1.2 The influence of corpus use on L2 learning and writing……………….247
  6.1.3 Collocation and L2 writing………………………………………………254
  6.1.4 Becoming L2 writers in EAP contexts………………………………….258
6.2 Pedagogical implications……………………………………………………….263
6.3 Limitations of the study………………………………………………………268
6.4 Recommendations for future research……………………………………….270

Bibliography…………………………………………………………………….273
Appendices...............................................................................................................................................284

A: Recruiting survey..........................................................................................................................284
B: Style Analysis Survey (SAS).......................................................................................................285
C: Strategy Inventory for Language Learning (SILL).................................................................289
D: Corpus search log form..............................................................................................................292
E: Survey questionnaire..................................................................................................................293
F: Recruitment letter.......................................................................................................................299
G: Informed consent form..............................................................................................................301
H: Excerpts of interview transcripts translated from Korean into English............302
I: Survey results...............................................................................................................................306
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Overview of case study participants</td>
<td>79</td>
</tr>
<tr>
<td>4.2</td>
<td>Overview of participants’ views on language learning and writing experiences</td>
<td>98</td>
</tr>
<tr>
<td>4.3</td>
<td>Results, Style Analysis Survey (SAS)</td>
<td>132</td>
</tr>
<tr>
<td>4.4</td>
<td>Results, Strategy Inventory for Language Learning (SILL)</td>
<td>133</td>
</tr>
<tr>
<td>5.1</td>
<td>Participants’ purpose of corpus use</td>
<td>145</td>
</tr>
<tr>
<td>5.2</td>
<td>Participants’ corpus search items</td>
<td>148</td>
</tr>
<tr>
<td>5.3</td>
<td>The amount of the participants’ writing assignments</td>
<td>158</td>
</tr>
<tr>
<td>5.4</td>
<td>Participants’ writing course grades</td>
<td>207</td>
</tr>
<tr>
<td>I.1</td>
<td>The class’ corpus use, advantages</td>
<td>306</td>
</tr>
<tr>
<td>I.2</td>
<td>The class’ corpus use, problems/difficulties</td>
<td>306</td>
</tr>
<tr>
<td>I.3</td>
<td>The class’ responses to corpus use in ESL writing instruction</td>
<td>307</td>
</tr>
<tr>
<td>I.4</td>
<td>The class’ overall evaluations of corpus use in L2 writing</td>
<td>307</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>140</td>
</tr>
<tr>
<td>5.2</td>
<td>141</td>
</tr>
<tr>
<td>5.3</td>
<td>142</td>
</tr>
<tr>
<td>5.4</td>
<td>154</td>
</tr>
<tr>
<td>5.5</td>
<td>194</td>
</tr>
<tr>
<td>5.6</td>
<td>196</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Background of the study

Corpus research is an emerging and ever-growing field in second language (L2) pedagogy. The conspicuous development of computer technology has popularized corpus linguistics in language research. Corpus linguistic research enables extensive text analysis and provides empirical language description based on large quantities of authentic texts. An increasing number of studies have recently applied the insights and methodologies of corpus linguistics to L2 pedagogy and research (e.g., Aston, 1995, 1997; Biber & Conrad, 2001; Conrad, 1999, 2000; Johns, 1994; McCarthy & Carter, 2001; Wichmann et al., 1997).

Put simply, corpus research analyzes a “corpus” by using a “concordance program” in order to identify patterns of language use (Conrad, 2000). A corpus is a large collection of natural texts, chosen to represent a language or a variety, which, nowadays, is almost always in a computer-readable form (Sinclair, 1991). Through concordance programs, corpora provide multiple examples of lexical or grammatical features in a variety of contexts and offer such textual information as word frequency counts and collocation patterns (see section 1.7 for an example). The compilation of
corpora has immensely expanded in number, size, and types (see Flowerdew, 2002, for an extensive overview of corpus development).

Corpus-based research shares several underlying principles. One fundamental assumption is that empirical data is needed for more accurate descriptions of language use. Language description, which is the basis of language teaching, has been grounded on human intuitions or “armchair reflections” (Johns, 1991) rather than on empirical observation. While such intuitions are often observed to be unreliable because of their subjective nature, the corpus-based approach has been regarded as a viable source to represent actual language use since it produces objective data about language use (Biber, 2001; Biber & Conrad, 2001; McCarthy, 2001; McEnery & Wilson, 2001; Mindt, 1996, 1997). Another essential principle of corpus research, from a linguistic perspective, is that language is presented in terms of a lexico-grammar (Halliday, 1992; Sinclair, 1991), which is distinguished from the traditional separation between vocabulary and grammar (Marco, 2000). That is, instead of looking at words in isolation, the corpus approach emphasizes the co-occurrence and typical contexts of words, i.e. “collocation.” In addition to collocation, frequency information of linguistic features is another important contribution that the corpus approach provides for language description.

In this view of language, a commonly accepted pedagogical premise is that “the most common words and their combinations should form the basis of instruction” (Jabbour, 2001, p.298). A great number of language educators have emphasized the usefulness of frequency and collocation information in order to prepare learners for the most common language use they are most likely to encounter (Biber et al., 1998; Biber & Reppen, 2002; Conrad, 1999, 2000; Jabbour 1997, 2001; Kennedy, 1987a, 1998; McCay,
1980; Mindt, 1996). In particular, corpus-based studies in L2 pedagogy have asserted the value of the corpus approach for teaching vocabulary and grammar, while corpus analysis is based on authentic language use and fosters inductive learning in the learning process. Some researchers even assert that the corpus approach will revolutionize language teaching (e.g., Conrad, 2000), although some scholars caution against extreme optimism regarding use of corpora in language pedagogy (see section 2.3 for detailed discussion about this issue).

Of special note is that corpus-based studies have concentrated on written production, although spoken corpora are increasing in number (Flowerdew, 2002). This is partly due to the ease of access to written texts as opposed to the difficulty in collecting speech samples, but also because text analysis is an indispensable methodology of corpus research. Accordingly, corpus research has revealed useful information about reading and writing, and, not surprisingly, corpus studies have often been associated with the teaching of L2 literacy skills. In this vein, Jabbour (2001) claims that “a corpus approach befits teaching second language reading and writing, since both activities are text oriented and make use of words and word combinations, or lexical patterns, within the confines of discourse” (p.294). On the other hand, it is also true that a very limited number of studies have addressed the use of corpora with respect to reading. A greater number of corpus studies have focused on how to use corpus technology in L2 writing research and pedagogy.

The study of the use of technology in L2 writing is by no means a recent topic of inquiry. As the use of the computer in academic contexts has increased, technology has ushered in an era of computer-assisted writing. While the bulk of computer-assisted
writing research centered on word processors until the mid-1990s (Bernhardt, Wojahn & Edwards, 1990; Haas, 1989; Hawisher, 1987; McAllister & Louth, 1988; Pennington, 1993a, 1993b), more recently other views of technology have become available for use in writing instruction (Ferris & Hedgcock, 1998). Most of the previous studies in computer-assisted writing research focused on the use of technology for interactions between learners and teachers or interactions among peers, electronic responses, and revisions through technology (e.g., Dalgarno, 2001; Tuzi, 2001). In other words, technology in writing research has been perceived mainly as a medium rather than a linguistic resource for writing. In this regard, corpus technology is distinctive compared to other kinds of technology used in composition research because it provides invaluable linguistic information to help students acquire common uses of lexis and grammar and ultimately improve their writing.

As a matter of fact, while current writing research has mainly addressed the so-called more global or discoursal aspects of writing, such as content and rhetoric, “a lack of grammatical or lexical accuracy can [still] be a major issue for L2 writers” (Ferris and Hedgcock, 1998, p.274). While L2 writers are well trained in developing their ideas, the linguistic domain is often a major challenge for many of them. It appears that grammar-focused instruction does not adequately prepare them for the linguistic demands of L2 writing. Therefore, L2 writing research and pedagogy must address this issue in order to help students become more competent writers. Here corpus technology is a promising tool to help L2 writers deal with linguistic problems in writing. In fact, Tribble (2002) asserts that the corpus approach is particularly useful for context knowledge (“knowledge of the social context in which the text will be read, and co-texts related to the writing task
in hand”) and language system knowledge (“knowledge of those aspects of the language system necessary for the completion of the task”) in writing. Yet, it is important to note that the corpus-based approach does not focus only on form at the expense of meaning. Tribble (1999) and Jabbour (2001) argue that the corpus approach connects form and function in the teaching of writing; it not only raises learners’ language awareness, but also contributes to an understanding of functions of linguistic features in context. For that reason, corpus technology might be a revolutionary tool for the teaching and learning of L2 writing. In sum, corpus technology appears to be a valuable linguistic resource to enhance L2 students’ writing experiences.

1.2 Statement of the problem

As the interest in using corpus technology in L2 writing pedagogy grows, researchers agree that the corpus information is particularly beneficial in materials development and syllabus design. A substantial number of corpus studies have been involved in developing corpus-informed syllabi, teaching materials, and classroom activities (e.g., Conrad, 1999, 2000; Flowerdew, 1998; Johns 1991, 1994; McCay, 1980; Thurstun & Candlin, 1998; Tribble, 1991). Those studies have emphasized that the corpus approach can not only enhance learners’ awareness of contextualized grammar, but it can also expand their understanding of the functions of words in context. Whereas early corpus research had an impact on development of dictionaries and grammar references, researchers have begun to look at academic written discourse, in combination with genre analysis, to inform English for Academic Purposes (EAP) materials (J.
Flowerdew, 2002) and “help students to develop competence as writers within specific academic domains” (Tribble, 2002, p.131).

While most corpus studies have mainly focused on materials development, there has been little empirical investigation into the use of corpus technology in classroom practice. As a result, although the previous studies have presented a general understanding of corpus use in L2 writing, we still lack insights into how corpus technology can be integrated into an actual writing classroom and how much it can contribute to the development of L2 writing skill. Evaluation is needed from students as well as teachers in order to judge the value of corpus use and to expand and/or qualify the use of corpus technology in writing pedagogy. Cook (1998), a well-known critic of corpus-driven language teaching, appropriately points out that corpus-based studies have yet to consider “students’ and teachers’ attitudes and expectations, the personal relationships between them, their own wishes, or the diversity of traditions from which they come” (p.58). In order to develop a fuller understanding of whether and how corpus technology contributes to students’ development into more competent writers, we must bring in multiple perspectives on the teaching and learning of corpus-assisted L2 writing.

Relatively few studies have examined students’ writing experiences in association with corpus use and their attitudes toward corpus use. Moreover, those studies are limited in terms of their scope and data collection methods. The studies have addressed student reactions to a corpus-based lesson (Sun, 2000), the importance of the training of students in the corpus approach for their own use (Turnbull & Burston, 1998), and the effectiveness of independent corpus investigations (Kennedy & Miceli, 2001; Fan & Xu, 2002). Notably, most of the studies have focused on teaching the corpus approach per se
rather than incorporating it into the writing process. In terms of data collection procedures, many of these studies conducted a one-time evaluation of students’ use of the corpus approach within a short time and provided limited qualitative insights (Fan & Xu, 2002; Sun, 2000), or else they studied a very small sample of participants with little use of corpora (Turnbull & Burston, 1998). In short, the previous studies did not fully illuminate students’ corpus use in L2 writing and its impact on their attitudes and writing in depth, thus resulting in a limited understanding of the role of corpus use in student writing development.

Furthermore, few studies have attempted to identify the profiles of learners using corpora in writing. In fact, many corpus studies have regarded learners as a monolithic group rather than as idiosyncratic individuals. Some research assumed differences in the effect of corpus use on language learning related to personal backgrounds, such as language proficiency, familiarity with the new approach, and learning styles (Johns, 1991, 1994; Stevens, 1991; Turnbull & Burston, 1998). They often assumed that intermediate and advanced learners benefit more from this approach than less proficient learners. However, researchers have not yet fully explored the students’ individual experiences, especially related to L2 learning and writing, and learning contexts that mediate corpus use and its influence on L2 writing. Given that teaching is extremely sensitive to the needs of different learners, thus making it hard to apply a universal principle across learners, the evaluation of the corpus approach in the individual learning and writing process is needed.

In summary, several gaps from the existing literature motivated this study. Most important of all, further research needs to be carried out to examine how corpus
technology affects students’ development of competence as L2 writers as well as their
writing experiences. The development of learner-specific descriptions in corpus use and
its evaluation would also be valuable. Also, there is little research on the teacher’s
integration of a corpus component into L2 writing instruction. In order to understand
how this new technology affects and contributes to students’ L2 writing, we need to look
at the phenomenon in depth in teaching and learning practice. Thus, while the existing
body of literature enhances our understanding of the usefulness of corpora in L2 writing
as composite parts, a triangulation of multiple perspectives from teachers and students is
needed, which in turn requires a triangulation of multiple methods.

1.3 Purpose of the study

The primary concern of this study was to gain a deeper understanding of the role
of corpus technology in L2 academic writing. In particular, it aimed to uncover the
changes in ESL students’ writing experiences associated with corpus use, and to
determine the extent to which the technology can facilitate the writing process. The
research site of the study was an ESL academic writing course, as tertiary-level ESL
students are often considered the most appropriate target group for the corpus approach
that promotes discovery learning (Turnbull & Burston, 1998). In order to develop more
complete descriptions of the phenomenon, it was considered desirable to combine
multidimensional perspectives from those involved in teaching and learning practice.
Therefore, while the focus of the study was on students’ use of corpus information as a
linguistic resource in their writing, this study also looked at the teaching side in the sense
that the teacher’s approach in using corpora affects the students’ use of the corpus and their perspectives on its use in writing.

Within the overriding purpose of the study, the subordinate objectives of the study were threefold: 1) to examine ESL students’ perspectives on using corpus technology in L2 writing, 2) to explore how corpus technology and in particular collocation affects their L2 learning and writing, and 3) to investigate any individual experiences and contextual factors mediating the influence of corpus technology on their L2 academic writing.

This study examined students’ uses of corpora while writing, their reactions to corpus use in writing instruction, and their evaluations of corpus experience in L2 writing. One of the key variables this study investigated was student attitudes toward writing and corpus use, and any changes in their attitudes as a result of corpus use. The evaluation of students’ attitudes is an important indicator of writing quality as well as writing experiences, because, as Pennington (1993a) put it, “attitudinal or social assessments [in individual students and in the writing class as a whole] directly measure the quality of students’ writing experience. Where this writing experience is positive, a beneficial effect has, in fact, already been achieved” (p.244).

In an attempt to explore in greater depth the influence of corpus use on student writing, this study examined the writing process associated with corpus use over time, investigating how corpus use affects the way students deal with linguistic issues in writing and the ways they approach L2 writing. Also, as collocation is an essential feature of the corpus approach, its role in L2 learning and writing was also explored through the students’ perspectives. Additionally, the study considered a variety of students’ individual experiences and learning contexts so as to deepen our understanding
of corpus use in ESL tertiary classrooms. The obtained learner profiles can help us identify its use in the classroom depending on learner characteristics.

In summary, this study aimed to provide insights into using corpus technology in L2 writing and suggest pedagogical implications for the teaching of L2 writing. With this information, writing teachers and L2 educators can adjust their classes to incorporate corpus technology appropriately in their classrooms.

1.4 Research questions

Given the primary purpose of the study, the following research questions were identified:

1. How do the students use corpus technology in L2 writing?
2. How does corpus use mediate the students’ understanding of language and L2 writing?
3. How does the native language influence the students’ L2 collocation uses? What role does collocation play in L2 learning and writing?

1.5 Overall research methodology

The main research site of the study was a graduate level ESL academic writing course at a large midwestern American university. In an attempt to investigate how the corpus is used in an actual classroom, the study was conducted in a natural classroom setting. The qualitative approach was chosen as a dominant paradigm to guide this study, as the study of the writing process, like this research, requires a qualitative perspective. The researcher attended every class, taking on a role as a participant-observer. The participant observation identified corpus-related classroom activities and related student
behaviors as they occurred. At the same time, six case studies were conducted to closely examine L2 students’ writing and corpus experiences in depth during and after the writing course. The case study methodology included interviews, recall protocols, think-aloud protocols, and document analysis. While this research was heavily based on the qualitative approach, it also used some quantitative data. A class survey was conducted to provide a general understanding of the class’ overall attitudes toward corpus use in L2 writing. This combination of the quantitative and qualitative aimed at obtaining a rich array of data for the study. All in all, triangulation of multiple methods and data sources provided more comprehensive and rigorous accounts of the topic under investigation.

1.6 Significance of the study

This study is of particular importance in its attempt at broadening the scope of computer-assisted writing research while using a triangulation of multiple methods. Although the existing literature provides a general understanding of computer-assisted composition, the impact corpus technology has had on L2 writing has not been the central focus of those studies. Therefore, the endeavor to better understand corpus-assisted writing makes this study unique in the existing body of literature.

In order to better understand the role of corpus technology in L2 writing, this study aimed to foreground the voices of individual students, while combining multiple data sources. In particular, few studies have closely examined students’ evaluations of corpus use as related to their writing processes. However, without looking at the whole process of writing from the perspectives of students themselves, we lack insights into what role corpus technology plays in students’ writing development. Furthermore, this
study is valuable from the methodological point of view. Little research has used the triangulation of methods in order to produce the rich descriptions intended in this study. Considering the complex nature of L2 writing and the recent development of corpus-assisted writing, the triangulation of multiple methods and sources can provide valuable insight into ESL students’ L2 academic writing as well as the influence of corpus technology on their L2 writing.

Consequently, this study has important implications for L2 writing pedagogy with respect to obtaining deeper understanding of the nature of corpus use in L2 writing and how it can contribute to the teaching of L2 academic writing. Based on the fuller understanding of the role of corpus technology in L2 writing, writing teachers and L2 educators can more effectively incorporate the corpus-based approach into the curriculum so as to enhance ESL students’ writing experiences.

1.7 Definitions of key terms

**Corpus linguistics**

Corpus linguistics is “the empirical study of language relying on computer-assisted techniques to analyze large, principled databases of naturally occurring language” (Conrad, 2000, p.548). One of the most distinctive characteristics of corpus linguistics is that it studies language as revealed in ‘real life’ language use, i.e., empirical data. As computer technology has developed immensely in its ability to store and process language, corpus linguistics is proliferating linguistic research and producing new facts about language. Like other disciplines of linguistics, corpus linguistics
provides descriptions and explanations of language, but it has “a tendency sometimes to focus on lexis and lexical grammar rather than pure syntax” (Kennedy, 1998, p.8).

**Corpus / Corpora**

A corpus generally means any collection of more than one text. But from the corpus linguistics viewpoint, it refers to a large collection of natural texts sampled to be representative of a variety or a genre of a language, which is now almost always in machine-readable form (Biber et al., 1998; McEnery & Wilson, 2001; Sinclair, 1991). Because a corpus is inevitably drawn from a sample of people using the language, the issue of representation is a crucial concern in corpus production. The purpose of the corpus production determines the compiling sources, size, and design of the corpus. In terms of size and design, corpora can be divided into two kinds: general corpora and specialized corpora (Flowerdew, 2002). General corpora are designed for general descriptive purposes of language, so they are often extensive databases of the texts (e.g., Bank of English and British National Corpus). In contrast, specialized corpora, often lesser in volume, are compiled from certain types of genres for specific research or teaching purposes.

**Concordance**

A concordance is an “exhaustive list for the occurrences of the word in context” (Biber et al., 1998, p.15), or a “display of words or simple grammatical items with their surrounding text” (Conrad, 1999, p.2). Sinclair (1991) states that “the concordance is at the center of corpus linguistics, because it gives access to many important language patterns in texts” (p.170). The most common format of the concordance is KWIC (Key
Word in Context) where the keyword is arranged in the middle of each line, with context on the left and right side of it. For example:

to be priests? It is hard to think of any reason why they wouldn’t be. Women’s skills do not work better. This is the biggest reason why it is worth bothering with this issue reason to cover them up. All the more reason for Greenpeace to expose them. We will of human rights. Let me give you one reason why the work of these researchers is so , Oxfam never actually loses out. One reason for rumors about supplies going stray is have to tell you that I am unhappy! The reason for this is that your administration is after, a health problem or other special reason, you may be able to train part time. You who just want to talk. Whatever the reason, the caller knows that The Samaritans will

Concordance program / Concordancer

A concordance program, also called a concordancer, is an essential tool for corpus linguistics to search for target words in a corpus and generate analyses of the words. The program not only produces concordances, but also provides a range of text analysis, such as frequency information and collocation patterns.

Collocation

Collocation is the patters of the co-occurrence of linguistic features in texts, that is, which words commonly appear together. A crucial notion of corpus linguistics is that looks at words in typical surrounding contexts of occurrence rather than in isolation. The term “collocation” is traced back to Firth, who proposed “the meaning of a word was much a matter of how the word combined textually with other words” (McCarthy, 2001, p.62). For example, McCarthy (2001) points out that the combination of “happy marriage” is often observed rather than that of “content marriage.”
**Frequency**

Frequency is the occurrence of linguistic features in texts. “Corpus linguistics makes it possible to describe the factors associated with the appropriate use of alternative structures when more than one structure is grammatically acceptable” (Conrad, 2000, p.554). Frequency is one of the crucial features of corpus linguistics that views language as a probabilistic system (Halliday, 1992). The word frequency list can be used as essential information for language teaching. For instance, Mindt (1996) explains that irregular verbs in English can be ordered by the frequencies of their occurrences based on corpora, and if learners acquire the most frequent ten words on the list, then it means they come to know 45% of English irregular verbs in use.

**Corpus technology**

In this study, corpus technology refers to computer technology used for corpus research that employs a corpus and a concordance program to produce concordances, frequency counts and collocation information of linguistic features and thus identifies patterns of language in use within the corpus.

**Corpus approach**

This term refers to the analytical approach that corpus linguistics uses to describe languages. It employs methodologies of corpus linguistics, which includes analysis of empirical data (computer corpora) and identifying typical patterns of language use, with a focus on lexis and collocation in language description. Thus, in this study, the term “the corpus approach” is used interchangeably with “the corpus linguistics approach.” Furthermore, the term is also used to refer to students’ use of corpus linguistic tools and resources, which entails corpus search and analysis.
**Natural text**

Natural text is a text from observation of language in actual use, as opposed to contrived and artificial constructions. The text is collected from language used for certain communicative functions, either in spoken or written form. For example, written corpora are drawn from magazine and newspaper articles, letters, memos, etc.

**Context**

Context can mean both the linguistic and nonlinguistic environment of any instance of language. More specifically, the former indicates a surrounding textual environment, which is often termed “co-text,” while the latter addresses the sociocultural background of any language activity. In this study, as Sinclair (1991) put it, “context can be assumed to mean normally the surrounding language, but not necessarily excluding the nonlinguistic environment” (p.171).

**Writing process**

The writing process is often associated with multiple stages of writing such as drafting, composing, revising, and editing (Zamel, 1983). It refers to a continuous process of writing activities that produce the final piece of writing rather than looking at the final product as an isolated outcome from the continuum. In this study, the writing process means not only a process-oriented approach to writing, but also any series of actions involved in the accomplishment of final written texts.

**1.8 Assumptions of the study**

One basic assumption is that the general corpus used in this study is large and representative enough to characterize and represent real language as it is used. In
addition, although the corpus is based in England, it is assumed that the corpus is appropriate for use in the American research site because it also includes texts from American English, and one can specify a search of American or British texts.

It is further assumed that there is a relationship between computer corpora and the writing process used by individual tertiary ESL students to develop their writing proficiency. This assumption allows the present research project to correlate normative data (i.e., corpus-based patterns of language use) with individual student samples of writing (i.e., academic ESL course writings).

This study also assumes that an investigation of ESL students’ experiences with writing is correlated with their English writing products. However, the exact nature of that correlation is not well defined. But it is very likely that studying students’ writing experiences will help us understand the individual processes that would result in particular products.

Another assumption this study makes is that knowledge is not discovered, but rather constructed through interactions between the researcher and the researched. As the present study was conducted mainly within a qualitative tradition, I did not attempt to maintain a distance from the participants, unlike traditional quantitative methods that view the researcher as a potential contaminant to the data, and to be separated out and minimized. Rather, the qualitative approach as applied to this study sought a reciprocal, dialogic relationship between “researcher” and “researched” (Wong, 1994) in order to obtain context-specific, grounded knowledge developed from the field-based and situated methodology. Likewise, qualitative inquiries do not aim to discover existing knowledge, but to construct new knowledge through interactions with participants.
CHAPTER 2

REVIEW OF THE LITERATURE

This chapter discusses relevant literature in the field of corpus studies and L2 writing research. Five areas of scholarship are reviewed: 1) Corpus linguistics and its view of language, 2) Collocation and L2 (writing) proficiency, 3) Controversies about using the corpus approach in language pedagogy, 4) Corpus studies in the development of classroom materials and activities, and 5) Research on students’ use of corpus technology in L2 academic writing. In order to provide some background of corpus studies, the first section, 2.1, begins with a brief overview of corpus linguistics and its view of language and language use. Section 2.2 focuses on collocation, an essential component of the corpus approach, and reviews the role of collocation in L2 learning and writing. Section 2.3 deals with the theoretical arguments about the use of the corpus approach in language pedagogy. Pros and cons about using corpora in language teaching are reviewed, focusing on the notion of authenticity and learner autonomy. Section 2.4 examines corpus studies that are mostly concerned with how to develop teaching materials and classroom activities. It addresses three areas: comparative analysis of textbooks and corpora, sentence-based corpus analysis, and genre-based corpus analysis. The last section, 2.5, then looks more specifically at corpus studies in L2 academic writing. In
particular, it focuses on the studies regarding students’ own use of corpus technology in L2 writing, beyond using corpus findings for materials development and syllabus design.

2.1 Corpus linguistics and its view of language

Corpus linguistics can be described as “the study of language based on examples of ‘real life’ language use” (McEnery & Wilson, 2001, p.1). While traditional linguistic analysis is mainly the study of the structure of language, i.e., what is theoretically possible in a language, corpus linguistics is concerned with the study of language use, i.e., how speakers and writers actually use the language (Biber et al., 1998). Underlying the corpus-based approach to language is the belief that an empirical investigation of authentic language use is a better source than native speakers’ intuitions or elicitation of artificial sentences for describing and explaining language. Sinclair (1997) emphasizes the value of corpus research that shows patterns of language in actual use. According to Sinclair, the first lesson from the corpus study is that “language cannot be invented, but only captured” (1997, p.31).

Corpus linguistics views language as a socially embedded phenomenon rather than an abstract system (McCarthy, 2001). Some argue that corpus linguistic perspectives are in line with the sociocultural theory on language, which views language essentially as a social semiotic tool and emphasizes language use through interaction in context rather than accumulation of language rules in one’s head (Lantolf, 2000). Tao (2001) claims that the corpus approach reveals how language routinely behaves in actual discourse, which matches a sociocultural linguistic perspective. McCarthy (2001) also notes the similarity between the corpus-based approach and sociocultural theory with an
emphasis on “the inseparability of language choices, social activity, the construction of relationships and the pursuit of goals” (p.88). In her review of sociocultural perspectives on language, Hall (2003) expands on a point made by Halliday, one of the pioneers of corpus studies. Hall explains that his view of language as a socially embedded activity is concerned with “the meanings of language forms in their systematic connections between the functions they play and their contexts of use” (p.25).

The goal of corpus analysis is to uncover typical patterns and establish a theory of language from the usual and typical patterns. Thus, corpus linguists are interested in “what frequently and typically occurs,” and accordingly “priority is given to describing the commonest uses of the commonest words” (Stubbs, 2001, p.151). From the corpus linguistic perspective that emphasizes frequent and typical patterns, language is conceived as a probability entity, rather than a series of deterministic rules (McCarthy, 2001). While quantitative data about frequencies comprises the essential part of corpus findings, corpus studies also search for functional and qualitative interpretation to explore systematic patterns and characteristics of language (Biber, 1988; Conrad, 1999).

As McEnery and Wilson (2001) indicate, the development of corpus linguistics has taken a few turns that can be characterized as three phases: early corpus linguistics, the Chomskyan revolution and its rejection of the corpus approach, and the revival of corpus linguistics as a promising methodology in modern linguistics.

According to McEnery and Wilson, linguistics prior to Chomsky was “entirely corpus-like” (2001, p.2). Its linguistic analysis was based primarily upon observed language use. The main fields of studies were biblical and literary studies, lexicography, grammatical studies, and language education studies (Kennedy, 1998). The early corpus
linguistics was the era of “pre-electronic corpora”: it did not use the computer and relied on human hands. Therefore, its analysis had limitations in data processing and thus was not considered to be accurate.

The early corpus approach was harshly criticized by Chomsky, who discredited the corpus as a source of information. According to McEnery and Wilson (2001), Chomsky’s criticism was based on three main points. First, Chomsky suggested that language should be modeled from competence rather than from performance because performance is an imperfect manifestation of competence. Therefore, for Chomsky, a corpus – which is performance data – can never be a source of information on which to model language. Secondly, corpora are partial, i.e., incomplete and skewed. Thus, Chomsky believed that a corpus, which consists of a finite set of partial examples, cannot possibly explain our knowledge of grammar that generates an infinite number of sentences. Thirdly, Chomsky noted what he saw as the absurdity of wandering in an enormous amount of untidy data, while we can easily determine the rules through introspection in the head. Chomsky’s linguistic theory became the dominant paradigm in linguistics in the 1960s. His rational theory of language was so influential that the corpus-based empirical methodology was rejected in linguistics.

However, many researchers, including McEnery and Wilson, presented eloquent counter-arguments. Sinclair, one of the leading corpus linguists, asserts that “the comprehensive study of language must be based on textual evidence [because one] does not study all of botany by making artificial flowers” (1991, p.6). Sinclair claims that the unpopularity of corpus-based language description was partly due to the difficulties in collecting data and analyzing typical examples. But, he continues, “the difficulties
should not be allowed to support the absurd notion that invented examples can actually represent the language better than real ones” (1991, p.5). Moreover, due to the marriage of the computer and corpus, corpus linguists are able to process huge amounts of language data in a systematic way, overcoming the limitation of early corpus linguistics. In fact, the term “corpus” is now almost always equated with “computer-readable corpus.” Stubbs (1993) also asserts that introspective data is not firsthand evidence: “since [the data is] selected to illustrate a theory, [it is] carefully considered and highly edited” (p.13). In addition, as McEnery and Wilson (2001) argue, introspective data is, in all likelihood, highly skewed as much as corpus data is. Pointing out that one’s intuition is not open to investigation by others, McEnery and Wilson present a strong case for using corpus data:

Chomsky’s criticism that we would never find certain sentences or constructions in a corpus overlooked an important point. If we do not find them, this is an interesting and important comment on their frequency [emphasis in the original] … Here we can see one example of data which is not susceptible to recovery via introspection, that is, frequency-based data (2001, p.15).

With the conspicuous development of computer technology, the corpus approach has become a prominent methodology in linguistic research. It is of note, however, McEnery and Wilson conclude that one important reason for the revival of corpus linguistics is to balance between introspective data and natural data in linguistic research. Likewise, many corpus linguists claim that the corpus and introspection-based approach should be seen as complementary, not as contradictory (e.g., Biber et al., 1998; Fillmore, 1992; Kennedy, 1998).
A new school of modern corpus linguistics started in the early 1960s. The school was originally formed in the British tradition of Firthian linguistics, which was field linguistics as opposed to the then dominant psycholinguistic approach (Stubbs, 1993). Neo-Firthians, such as Halliday, Sinclair, and Hoey, are one of the most influential groups that laid the foundation to modern corpus research. Neo-Firthians reject Chomsky’s distinction between competence and performance; rather, they aim to model language paradigmatically by studying the samples of language in use.

More importantly, neo-Firthians are interested in lexis as an independent area of study of language (McCarthy, 2001). Although some regard corpus linguistics as a methodology that can be used in any branch of linguistics, rather than one that confines itself to a specific area of linguistics (Kennedy, 1998; McEnery & Wilson, 2001), it is true that the corpus linguistic approach emphasizes the role of lexis in the description of language. This is in contrast with traditional linguistic analysis, which is mainly concerned with syntax. In traditional linguistics, lexis, which is full of irregularities, is often viewed as secondary as opposed to systematic syntactic rules (Francis & Sinclair, 1994). In contrast, the corpus-based approach to language places a great emphasis on lexis, and thus the corpus approach is often referred to as a “lexical approach.” Mainly due to the growth of corpus linguistics, recently, lexicography has been brought to the front in applied linguistic research, “offering the possibility of a socially sensitive theory of lexis” (McCarthy, 2001, p.62).

Furthermore, from the corpus approach to language description, “lexical patterning is seen as the key to grammatical description” (Owen, 1993, p.167). Sinclair is one of the leading figures who promoted a more lexically-based approach to linguistic
The Collins Cobuild English project, led by Sinclair, aims to “specify all major lexical items in terms of their syntactic environments, and all grammatical structures in terms of their key lexis and phraseology” (Francis & Sinclair, 1994, p.199). As such, perhaps not surprisingly, Sinclair rejects the separation between form and meaning. For him, “there is ultimately no distinction between form and meaning” (1991, p.4).

Likewise, the corpus approach to language states that lexis and grammar interact in systematic ways, so it looks for the interdependent relationship between the two. Halliday is one of the earliest proponents to relate the study of lexis and grammar. For Halliday, grammar and vocabulary are not two separate things, but one phenomenon, which is in a continuum (1991; 1992). He thus proposed the combination of lexis and grammar, “which is referred to in everyday speech as the ‘wording’; technically it is lexicogrammar” (1992, p.63).

‘lexis’ and ‘grammar’ are names of complementary perspectives, like the synoptic and dynamic perspectives on a semiotic process, or wave and particle as complementary theories of light, each explaining different aspects of a single complex phenomenon. Given this concept of lexicogrammar, it does not make sense to condone relative frequency in lexis but deny its validity in grammar (1991, p.32).

While both Sinclair and Halliday argue for the meeting of lexis and grammar, their starting points are somewhat different; if Sinclair starts from the end of lexis on the continuum, Halliday starts from the end of grammar on the continuum, although they finally converge. In other words, while Sinclair’s approach is mainly driven by focus on lexis, Halliday’s use of corpus analysis is more anchored in the study of grammar. This can also be seen in that Halliday views lexis as the “most delicate grammar” (Halliday,
Halliday (1991) pointed out that while there is little objection to using probability profiles of words (e.g., *go* is more frequent than *walk*, which is in turn more frequent than *stroll*), there is resistance to the use of frequencies of patterns in the grammar. Halliday, however, perceives grammar as a probabilistic system as well. He proposes the notion of a probabilistic grammar in which “the linguistic system is inherently probabilistic, and that frequency in text is instantiation of probability in the grammar” (p.31). For Halliday, corpus findings are important because they can be utilized to build up the probability profiles of grammatical systems. But he points out that the usage of the probabilities is not to predict single instances, but to interpret the general pattern of language system. Halliday asserts that “the transformation of instance [“text”] into system [“language”] can be observed only through the technology of the corpus, which allows us to accumulate instances and monitor the diachronic variation in their patterns of frequency” (1991, p.34). Thus, Halliday emphasizes the use of the corpus as a source to develop a linguistic theory.

Although the use of corpora in language research is now an established approach (Halliday, 1991), there are still concerns about corpus production in terms of its representation of language. Because a corpus is a drawn sample aimed to characterize a language, the size as well as representativeness is a critical issue in its design. Sinclair (1991) points to the significance of representativeness in corpus design by succinctly stating that “the results are only as good as the corpus” (p.13). The representativeness of the corpus was of a great concern, particularly in the early ages of corpus studies. In fact, Chomsky’s attack on corpus data was largely based on the issue of balance and representitiveness. Yet, with a growing recognition of the importance of
representativeness in corpus design and the enormous advancement of computer capacity, the concerns seem to have diminished. A variety of corpora (e.g., International corpora, learner corpora, and multilingual corpora) are now being produced under more principled design. Nonetheless, this issue should still be taken seriously in the production and use of the corpus. McEnery and Wilson (2001) emphasize that when interpreting corpus findings, corpus linguists should always bear in mind that the corpus is a sample from a group of people using the language. The usefulness of corpora, then, depends on how one produces and interprets the results of corpus analysis. McCarthy (2001) also asserts that corpus technology is far from being neutral and thus should be understood with the ideological issues. In particular, a greater level of concern about using corpus technology has been raised from the pedagogical perspectives as an increasing number of studies have applied the corpus approach to language teaching. Controversies about using corpora in language pedagogy are reviewed in detail in Section 2.3.

2.2 Collocation and L2 (writing) proficiency

Sinclair (1991) argues that information about collocation is one of the most useful assets that corpus research can provide for language description. Collocation refers to strings of words that conventionally go together, which can be more easily understood as “wording” or “word combinations.” The term ‘collocation’ was first used by Firth in its modern linguistic sense, but it was popularized by Sinclair later (McEnery & Wilson, 2001; Partington, 1998). As Sinclair (1991) observes, collocation is originally confined to lexical association patterns, but it often expands to include the association with grammatical items. Many scholars claim that collocation is at the heart of lexicogrammar
that looks at words in lexical as well as grammatical surrounding contexts of occurrence. Kennedy (1998) sees the concept of collocation as the place “where grammar and lexis meet in the phrase” (p.289). Aijmer and Altenberg (1991) also assert that “collocations … represent the intersection of lexicon and grammar, an area which can be fruitfully studied in corpora” (p.4).

Whereas Chomskyan linguistics views the irregularity of collocation as a challenge to the rule-governed generative grammar, many scholars claim that human beings’ language use is not so much creative or generative and that the conventional and idiomatic use of language comprises an integral part of language (Hopper, 1998; Howarth, 1996; Pawley & Syder, 1983). In their groundbreaking article “Two puzzles for linguistic theory: nativelike selection and nativelike fluency,” Pawley and Syder (1983) claimed that nativelike usage is much more restricted and predictable than is often assumed. They presented alternative explanations about the idiomatic use of language, which was not clearly explained by the then dominant generative grammar approach to language. In order to solve the puzzle of native speakers’ ability to select natural and conventional usage among the wide possibility of grammatically correct sentences, they drew upon a ‘lexical sentence stem,’ “a unit of clause length or longer whose grammatical form and lexical content is wholly or largely fixed” (p.191). According to them, one of the key factors in the fixed word combinations or “form-meaning pairings” is social recognition or institutionalization. Importantly, they argued, using idiomatic phrases reduces the work of information retrieval and processing and allows speakers/listeners to engage in other discourse-level communication skills.
Pawley and Syder’s work is congruent with the corpus linguistics approach to language. From the corpus linguistics viewpoint, native speakers’ language use is not so much inventive as conventional, as is evident in fixed collocational patterns (Hill, 2000; Howarth, 1998; Partington, 1998). The collocational field has recently gained great attention from those interested in L2 acquisition and pedagogy. Hill (2000), who emphasizes the importance of collocational knowledge in L2 pedagogy, addresses a fundamental question of what it means to know a language. He claims that “[students] do not really ‘know’ or ‘own’ a word unless they also know how that word is used [emphasis in original], which means knowing something about its collocational field” (p.60). That is, knowing a language really means knowing a variety of collocation patterns and learning “word grammar.” Consequently, this view of language calls for a change in the traditional approach to teaching a second language. It asks for the combination of lexis and grammar rather than teaching grammatical rules and vocabulary separately.

Like Pawley and Syder, Hill (2000) argues that the use of fixed expressions can facilitate naturalness, fluency, and effectiveness in language use. Further, according to Hill, the barrier for intermediate students to move up to the advanced level is the lack of collocational knowledge. As he puts it:

**Spending a lot of class time on traditional EFL grammar condemns learners to remaining on the intermediate plateau** [emphasis in original]. Helping learners to become ‘advanced’ needs a huge injection of lexis. It is lexis in general, and collocational competence in particular, which allows students to read more widely, understand more quickly, and speak more fluently (2000, p.68).
Collocational knowledge has become widely recognized as an important component of L2 proficiency (Hill, 2000; Howarth, 1998; Lewis, 2000). One of the most difficult challenges for language learners is to acquire the natural, idiomatic word combinations that are commonly used in the language. While there is still a lack of a clear understanding of how L2 learners (even with sufficient grammatical knowledge) acquire native-like fluency that is not fully explained by rule-based formal system, the corpus linguistics approach can contribute to solving a chronic puzzle in L2 pedagogy: acquisition of conventional use and native fluency.


Despite their uses of different terms, the researchers seemed to agree that collocation should be understood as a continuum according to their restrictions of word combinations (i.e., from free combination to restricted combination). Howarth (1996,
1998) and Nesselhauf (2003) presented the most explicit continuum model. Both of them focused on “verb-noun combinations” and presented three major classifications of the collocation: free combinations (both verb and noun are used unrestrictedly, e.g., *want a car*), restricted collocations (the verb is used restrictedly and combined with certain nouns, e.g., *take a picture*), and idioms (both verb and noun are used restrictedly and often seen as a chunk, e.g., *foot the bill*). They went on to point out that traditional English education has primarily focused on the idioms at the expense of “collocations that are much less fixed in form than idioms and [thus] potential problems for learners” (Howarth, 1996, p.1). Then, as Farghal and Obiedat (1995) noted, what learners need first is “the awareness of collocational restrictions in English” (p.327). Further, Howarth (1996, 1998) emphasized the importance of the awareness in production rather than in comprehension:

The difference can be seen between problems of encoding and decoding. *Foot the bill* might pose problems of decoding, but would be avoided by most learners in production. Problems of encoding might arise more often at the free end of the spectrum, where learners are unaware of the arbitrary way in which restrictions operate (1996, p.44).

In particular, Howarth (1998) argued that “collocations can be considered most centrally involved in the process of composition at clause level, therefore potentially sensitive indicators of learners’ acquisition” (p.26). Many researchers agreed that collocation plays a particularly important role in production rather than in comprehension (Bahns & Eldaw, 1993; Gabry-_Biskup, 1990; Howarth, 1996,1998; Nesselhauf, 2003). As Gabry-_Biskup (1990) put it, “collocations do not cause a problem of perception (understanding) but that of production (p.35).”
A body of research based on this framework exists in order to investigate learners’ use of collocation in L2 writing. Although they adopted different foci and tasks for the study, most of them focused mainly on quantitative analysis of students’ one-time collocation use or final products of writing. Those studies can be categorized into three areas: 1) test on students’ collocation knowledge, 2) error analysis in students’ actual writing, and 3) comparison of native speakers’ writing with learners’ writing. The first area mostly used a cloze test or translation task to test learners’ L2 collocation knowledge, while focusing on a smaller unit of clause level production (Bahns & Eldaw, 1993; Farghal and Obiedat, 1995; Gabry_-Biskup, 1990; Hussein, 1990). The second area examined learners’ writing that was produced through composition tasks and developed error analysis in their collocational uses (Dechert & Lennon, 1989; Man-lai, Pui-yiu, & Chau-ping, 1994; Nesselhauf, 2003). Thus, the studies in the area often focused on the impact of students’ native language on their L2 collocation production and mostly presented the result of L1 interference. The last area often adopted corpus analysis to compare production data of native speakers with that of learners’ to identify differences in the use of collocation (Granger, 2001; Howarth, 1996, 1998).

The predominant quantitative analysis of students’ collocation uses leaves other areas unexplored. Few studies have examined L2 students’ collocation learning and writing processes, and the relationship between the two. In particular, little research has investigated students’ own use of corpus technology and its impact on their collocational knowledge and L2 writing. In order to fully understand how the corpus approach contributes to learners’ acquisition of collocation and L2 writing proficiency, however, we need to look at their writing processes as well as products. Further investigation is
needed to explore not only ESL students’ writing experiences associated with corpus use, but also how corpus technology can facilitate their collocational competence.

2.3 Controversies about using the corpus approach in language pedagogy

While it is generally accepted that corpus linguistics has made contribution to descriptive linguistics, more controversies have been raised as to the usefulness of corpora and the extent to which corpus findings should influence language teaching. This section discusses the controversies about using the corpus approach in language teaching and learning. First it reviews arguments in favor of the corpus approach in language learning: how the basic premises of the corpus approach can help students with their L2 learning. Four basic tenets of corpus studies are reviewed: the focus on lexis, the use of frequencies of words, the connection of form and function, and the emphasis on discovery learning. The arguments against using the corpus approach in language teaching then follow. The cons mainly address three aspects: extrapolation from linguistics to pedagogy, authenticity of language data, and learner autonomy in the learning process. Several counter-arguments to the cons are also presented.

Pros

First of all, the corpus approach has brought to light the importance of lexis in language learning. In the past few decades, mainly due to Chomskyan grammar, vocabulary has not been considered as central to language teaching, while an over-emphasis is placed on the structure of language. In contrast, as the corpus approach places lexis at the center of language description, its approach has emphasized the significance of lexis in language learning. Willis (1990) goes on to argue for a lexical
syllabus where instruction is constructed around vocabulary in language teaching. He contends that identifying the most frequent words by use of the corpus approach also enables learners to identify the most frequent notions and functions. Francis and Sinclair (1994) emphasize the importance of lexis in language teaching:

There is little point in presenting learners with syntactic structures – how groups and clauses are built up – and then presenting lexis separately and haphazardly as a resource for slotting into these structures. In other words, we should not burden learners with vast amounts of syntactic information on the one hand and lexical (‘vocabulary’) information on the other, which they then have to match according to principles which are not naturally available to them as non-native speakers. Instead, teachers can present the structures and their lexis at the same time, either in the form of lists, or by means of concordances if the teaching situation allows this (p. 200).

In addition, the corpus-based perspective states that a lexically based grammar should be emphasized in language teaching. In other words, the approach assumes that “vocabulary should lie at the heart of the teaching process, impinging on, or even guiding the teaching of all areas of language, including grammar” (Owen, 1993, p.171).

Furthermore, the corpus approach rejects the traditional approach to vocabulary teaching, over-emphasizing “single words out of context” (McCarthy, 2001, p.63). The traditional perspective often assumes that words have inherent meanings and thus can be taught in isolation from other words and syntactic structures. In contrast, corpus linguistics believes that “words did not have inherent meanings, but depended on their environments to select or at least confirm their meaning” (Francis & Sinclair, 1994, p.192). The corpus approach then looks for collocational meaning in vocabulary. This approach is particularly helpful for teaching learners the combination of words and the use of words in different contexts (Biber et al., 1998). Jabbour asserts that “the use of
collocations in context is essential for selecting and sequencing the teaching units” (2001, p.251).

The second emphasis of the corpus approach is on the use of frequency in language teaching. A great number of researchers have stressed the usefulness of frequency information of linguistic items in language teaching. For example, Kennedy (1987a; 1998) claims that if selection is unavoidable due to limited class time, the focus should be on the most frequent words and common patterns. According to Kennedy, “words are not useful because they are frequent, but frequent because they are useful” (1987a, p.283). He asserts that corpus studies help curriculum designers and teachers to determine “which language items and processes are most likely to be encountered by language users, and which therefore may deserve more investment of time in instruction” (1998, p.281). Conrad (2000) also makes the same point – a frequency-based syllabus, which is based on statistically significant data, enables teachers to focus on the most useful items for learners. The frequency-based syllabus can provide students with a number of representative examples that offers them a rich experience of language in use. It is important that students who are taught with this approach can more successfully relate their learning in the classroom to real language use outside class (Little, 1997).

The underlying principle of the frequency-based teaching is that when language is seen as a probabilistic system that requires a choice among items, the frequency information by corpus studies can help learners choose the most common ways of language use. However, it is important to note that the corpus approach does not state that everything is a matter of choice in language use. Carter (1998) addresses the misconception about the use of frequency in language teaching, particularly grammar.
It is, of course, misleading to suggest to learners of English that grammar is simply a matter of choices. Grammatical rules exist; they have been extensively codified, and form the core of the structure of (both spoken and written) language. … Within a central core, choices are not possible (p.51).

This comment reflects the middle ground view that values combination of rule-based and corpus-based approaches to language description, while seeing lexis as a crucial element of language teaching and learning, not as subsidiary fillers to grammar.

Another benefit of the corpus approach is that corpus use connects form and function in language learning. Johns (1994) claims that while communicative language teaching has generally been associated with a focus on language function rather than on language form, corpus-based authentic materials place an emphasis on form as well as on function. Jabbour (2001) makes the same point. Whereas the audiolingual method and communicative language teaching focus on either form or meaning, respectively, Jabbour notes that the corpus approach generally focuses on both forms and functions in teaching reading and writing. In an earlier work, Jabbour (1997) claims that the text-based corpus approach raises learners’ language awareness, which in turn contributes to their understanding of functions of linguistic items in a variety of contexts. In other words, class activities based on corpus-based text analysis provide learners with a large number of concrete examples for matching form and meaning, with a focus on both. Tribble (1999) also argues that the corpus approach, in combination with genre analysis, provides learners with opportunities to identify forms as well as content that are appropriate to a given genre.
Lastly, but perhaps most importantly, the corpus approach (particularly concordancing) places a great emphasis on students’ inductive learning in language pedagogy. According to Johns (1991), learners who use corpus data are able to discover the rules of the language inductively for themselves, which is opposed to traditional rule-based learning. He proposes a data-driven approach to learning where learners are viewed even as researchers.

Research is too serious to be left to the researchers: that the language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data – hence the term “data-driven learning” (DDL) to describe the approach (p.2).

Some also emphasize that learners become “independent analysts” (Odlin, 2001) by manipulating and analyzing corpus data. In so doing, learners make their own discoveries of language while testing, confirming, and generating the rules. Thurstun and Candlin (1998) assert that concordancing is one of the best ways to help students to make direct discoveries about language, rather than giving ready-made rules to be memorized. In this sense, the corpus approach – which is a student-centered learning approach – is believed to contribute to the development of learner autonomy.

**Cons**

Arguments against the use of corpus data in language pedagogy are mainly proposed by Cook and Widdowson. Cook and Widdowson raised questions about transferring the corpus linguistic view directly to the pedagogical application. Warning against “the leap from linguistics to pedagogy,” Cook (1998) argues that linguistics insights do not necessarily revolutionize teaching. Widdowson (1991) also claims that
the improved language descriptions of corpus linguistics do not necessarily provide a better basis for language teaching. Although he agrees that “[corpus findings] reveal a reality about language usage which was hitherto not evident to its users” (2000, p.6), he raises concerns about using the descriptive approach of corpus linguistics in language pedagogy. Widdowson criticizes the attempt to transfer corpus studies to pedagogy as “linguistics applied” rather than “applied linguistics” (2000). Rather, he argues for the prescriptive approach to language based on the criteria of contextual applicability, teachability, and classroom needs. Owen (1993, 1996) argues for combining corpus data and teachers’ intuitions in language teaching, rather than giving precedence to the observed data:

> trusting the corpus data to the exclusion of one’s intuition about what is possible in the language may have been a necessary antidote to hidebound convention in linguistics and language teaching. … The grammarian and the language teacher need the corpus as servant, not as master (Owen, 1993, p.185)

A more severe criticism toward the corpus approach comes from considering different contexts of learning. Cook (1998) attacks the abuses of corpus studies in language teaching, claiming that the collection of native speaker’s uses of the language is not necessarily beneficial to students in a different context. According to Cook, “corpus-driven language teaching always risks stressing what is actually done at the expense of what is appropriate in a particular context” (1998, p.60). Widdowson makes the same point that the real language can hardly be authentic in contexts of students’ learning. Widdowson (1993; 1994; 1996; 1998) disagrees with the general idea of authentic materials. He argues that any text brought from real communication is de-authenticated
in the classroom context and thus it cannot possibly be authentic for foreign language learners who study the text. Consequently, he contends, emphasis should be placed on contextual conditions instead. In other words, classroom materials should be recreated to be meaningful in the contexts of learning. Furthermore, for Widdowson, the notion of “context” in the sense of corpus linguistics does not address contextual factors in sociolinguistic or sociocultural sense. Rather, it only means “co-textual occurrence” (2001, p.536). Hence, Widdowson argue, “it is problematic to assign pragmatic significance directly to textual features” (2001, p.537).

Another strong counterargument to use corpus work in teaching is based on the proposition that authentic materials negate learner autonomy. Widdowson (1996) asserts that “authentic language is, in principle, incompatible with autonomous language teaching” (p.68). For Widdowson, authenticity is based on native speaker’s authority while autonomy is dependent on non-native speaker’s authority. Additionally, he continues, authenticity places a value on the goal of learning while autonomy values the process of learning. After all, Widdowson (1994) claims that the notion of authenticity gives primacy to native speaker use of the language as the proper language for learning and further privileges native speaker teachers in teaching.

On the other hand, many researchers provide counter-arguments to the criticism of the corpus approach. McCarthy and Carter (2001) agree that authenticity is not necessarily guaranteed by importing real language data into the classroom, and, therefore, contextual consideration is required in order to apply the corpus studies to classroom teaching. Nonetheless, they believe that the language descriptions of corpus studies can still offer substantial benefits for language instruction if well applied. They claim that the
pedagogical process should be informed, not driven or controlled, by the corpus to lead students to experience actual language use. McCarthy (2001) points out that Widdowson is too negative about the use of corpus linguistic findings and methodologies in language teaching by only drawing upon a “corpus-based” and “corpus-driven” approach, failing to explore an alternative “corpus-informed” approach. In contrast to Widdowson’s interpretation of “context” in corpus linguistics as a concept limited to immediate textual surroundings, McCarthy (2001) also argues that corpora are definitely social artifacts whose production includes a consideration of social contexts. He remarks that “in the real world of real applied corpus linguists, some of Widdowson’s fears seem far removed from actual practice” (p.129).

Furthermore, many scholars do not view the notion of authenticity as being opposed to the notion of learner autonomy. On the contrary, they claim that learners can develop autonomy in their language learning based on authentic materials. According to Little (1997), learners who are exposed to authentic texts can rapidly develop confidence in their actual use of the language and can thus make the easy transition from language learning to language use. For that reason, he claims that authentic materials contribute to the development of learner autonomy. Aston (1997) believes that corpus data can provide students with self-access learning experiences by exploiting available resources for themselves and acquiring keys to future learning outside the classroom. Carter, Hughes, and McCarthy (1998) argue that while teachers and learners can decide whether to use the corpus approach, learners should not be “disempowered” by the unavailability of the options:
Teachers and learners can always choose not to learn those areas of language where rules are more probabilistic than determinate but that they have no choice at all if such options are not made available. Learners should not be disempowered and syllabuses should not be deliberately impoverished (p.85).

The most advanced discussion of the issue is given by Gavioli and Aston (2001), who achieved a resolution to the controversy between authenticity and learner autonomy (Seidlehofer, 2003). They gave an example of how a learner solved his problem in writing by consulting corpus data and constructing language patterns. They explained that the learner, as a discourse participant as well as an observer of the real texts, constructed a model for his purpose by using corpus data, thereby authenticating the real texts in his own discourse rather than simply imitating them. Gavioli and Aston conclude that in the corpus approach, learners “authenticate discourse independently and collectively, adding to the reality of the corpus the reality of their own experience of it” (p.244). In this context, the two notions – authenticity and learner autonomy – are seen as reconciled rather than opposed.

In summary, as Aston (1995) indicates, the critics of the corpus approach may warn against the extreme optimism toward the corpus approach. According to Aston, Widdowson’s concern should not be interpreted to mean that the findings of corpus linguistics are necessarily inapplicable to language teaching. Rather, Widdowson cautions against an uncritical use of corpus data and places an emphasis on principles of how to use the descriptions of language for teaching. In fact, many corpus researchers themselves ask for a cautious adaptation of the corpus approach to language teaching. Kennedy (1998) states that “[concordancing] is not a language teaching methodology nor a panacea but one among many techniques or aids which may be used to facilitate
learning for some learners” (pp.293-294). Carter (1998) showed a strategy of creating teaching materials by taking “a middle ground between authentic and concocted data” (p.52), while modeling on corpus data. Tribble’s (1999) comment on corpus-based teaching is also worth noting here: “it is import [sic] that [the corpus approach] be seen as complementing rather than replacing other, more familiar, classroom activities” (p.279). It may be impossible or ineffective to replace the conventional way of teaching with the new corpus approach. Hence, in order to lead to a successful use of the corpus approach, attention should be paid to the integration of the approach with the existing curriculum in a particular pedagogical context.

The remaining part of this chapter examines previous corpus studies in language teaching. As Partington (1998) indicates, corpora in language pedagogy can be used in two different approaches: 1) the development of teaching materials or teacher’s use of corpora for syllabus design, and 2) the exposure of students to corpora for their own use. The next section examines the first approach, while the last section reviews the studies regarding students’ own use of corpora.

2.4 Corpus studies in the development of classroom materials and activities

The views of language and language use are important for language teaching because they provide the baseline for materials development and classroom teaching. In particular, the development of teaching materials requires continuous judgments about language use in order to present the language as it is used. Those judgments have usually been based on tradition, intuition, or anecdotal evidence of language use, rather than on empirical evidence (Biber, 2001; Biber & Conrad, 2001; Mindt, 1996, 1997). Along with
the growing popularity of the corpus approach in language description, many researchers have recently argued for using the insights and findings of corpus research in the development of materials and syllabi in L2 pedagogy. Gavioli and Aston (2001) claim that empirically-based corpus data can help material developers and teachers test oft-wrong intuitions and focus on frequent uses of the language, thereby resulting in better-informed materials and syllabi in language teaching. Conrad (1999) makes the same argument – “corpus-based study can lead to more principled classroom materials and activities” (p.1). Biber et al. (1998) also identify three ways that corpus findings can be used in language teaching: the development of curriculum in specific domains, the development of teaching materials, and the design of classroom activities.

In particular, many L2 educators have argued that teaching materials should be informed by corpus studies (Carter et al., 1998; McCarthy & Carter, 2001). Considering that corpus linguistics provides new insights into language description, it is no surprise that its findings can offer valuable information for materials development (Conrad, 1999, 2000; Thurstun & Candlin, 1998). In fact, the development of materials has been regarded as the area that can benefit most from incorporating the corpus approach into language teaching.

Admittedly, English teaching materials, particularly textbooks, have long provided artificial examples of language use due to the strong prevalence of pedagogically contrived texts over authentic texts (Johns, 1994). However, there has been a growing interest in the use of authentic materials as communicative language teaching approaches have been emphasized in language pedagogy. The corpus has been presented as a valuable resource to provide authentic materials. And many scholars claim
that “the findings of corpus-based investigations can be used to inform the presentations in textbooks for ESL students” (Biber et al., 1998, p.80). Many studies have attempted to employ the actual corpus data in designing teaching materials so that the textbooks can reflect more accurate language use.

Several researchers have used corpus findings to critically examine English textbooks (Holmes, 1988; Kennedy, 1987a, 1987b; Ljung, 1990; Mindt, 1996, 1997). These studies used similar methods. They compared English textbooks with major (general) corpora to investigate whether the textbooks reflect actual language use as it is shown in the corpora. Their focus was on the examination of certain lexical or grammatical items of interest in the ESL/EFL textbooks. They analyzed the frequency and the treatment of those items in both sample textbooks and standard corpora of English. From their results, most of the studies discovered considerable differences in the presentations of the items between the textbooks and the corpora. Some textbooks emphasized the items that were found less frequently in real language use, while paying less attention to the items that were actually more often used. The researchers conclude that the findings of corpus-based studies should be incorporated into the textbooks in order to provide students with authentic language use.

Those comparative studies of textbook and corpus English are remarkable for their early attempts to apply corpus insights to materials development and broaden the scope of language teaching. Nonetheless, the studies were limited in their use of corpora to only checking the usage of a certain item rather than making use of the corpora to generate corpus-based materials.
With the development of various kinds of corpora, there has been another line of research that has used the corpora more extensively. Importantly, most of the corpus studies have focused on writing pedagogy. The studies can be divided into two areas: sentence-based corpus analysis and genre-based corpus analysis. Whereas corpus-based studies in the 1980s and early 1990s largely concentrated on linguistic analysis at the sentence level, corpus research has now expanded to embrace a broader genre-based analysis at the discourse level (Flowerdew, 2002). It is important to remember, though, that the studies share one common characteristic of corpus research, regardless of their use of a sentence or a genre as a unit of analysis. That is, they all look for lexico-grammatical patterning of texts, which is a central concept of any corpus research. In addition, they both aim to enhance materials development and syllabus design based on the insights and findings of corpus analysis (Conrad, 1999; Gavioli and Aston, 2001). In other words, their common objective is to employ more accurate descriptions of language use by corpus analysis to create teaching materials and activities with a particular focus on writing skills.

The first area, sentence-based corpus analysis, focuses on linguistic patterns of a text at the sentence level, while mainly drawing upon general corpora for analysis. Thus it investigates linguistic characteristics across genres and attempts to obtain general descriptions of language. Also, this line of research is largely concerned with a micro level of analysis, such as articles, prepositions, and certain words (Johns 1991, 1994; McCay, 1980; Thurstun & Candlin, 1998; Tribble, 1991). However, although the analysis is at a micro level, it aims not only to enhance students’ awareness of contextualized grammar, but also to expand their understanding of functions of words in
a variety of contexts. The studies of this area often demonstrate a strategy of designing corpus/concordance-based materials and tasks that can be used in the classroom. In so doing, they present possibilities of using corpus approaches in L2 instruction.

On the other hand, the second area, genre-based corpus analysis, has recently developed genre-based approaches into the corpus analysis of texts. The studies of this area look for identification of lexico-grammatical patterns of a particular genre rather than general descriptions across the genres. In other words, this area of research aims to illuminate idiosyncratic linguistic characteristics of a particular genre. Biber’s (1988) study is seminal in this area.

Biber conducted an exhaustive analysis of spoken and written genres by use of large English corpora. His study was built on the idea that a single dimension cannot explain textual variation among genres and that a multi-dimensional approach is necessary. Biber thus developed a “multi-feature/multi-dimensional” model of which methodology embraces both quantitative and qualitative analysis. In particular, his analysis was based on patterns of collocation as well as frequency counts of linguistic features. The main assumption of his analysis was “particular sets of linguistic features co-occur frequently in texts because they are serving some common communicative function in those texts” (p.13). As a result, he identified linguistic co-occurrence patterns that were grouped into factors, which showed underlying functional dimensions. His study contributed to establishing a linguistic basis for stylistic variation among genres.

Since Biber’s attempt, a growing number of studies have used corpus techniques in combination with genre analysis. Since corpus methodology and genre analysis are both based on text analysis, many researchers have argued for the usefulness of
integrating the two approaches (e.g., Gledhill, 2000). The emphasis of this line of research is on linguistic specification of a given type of discourse by use of corpus analysis. It is, then, no surprise that the studies usually employ specialized corpora, which are compiled from texts of the target genre. In particular, those studies have argued that the teaching of academic writing needs to focus on the most frequent linguistic and rhetorical features that are specific within a discipline so that students can become aware of the function of the common collocational frameworks of the target genre and thus finally improve their writing skills (Gledhill, 2000; Hyland, 2002; Jabbour, 1997, 2001; Tribble, 1999, 2002; March, 2000). It is worth noticing here that the studies often aim to provide insights into the teaching of a particular genre, which is commonly associated with the field of EAP/ESP. Consequently, the best use of those studies seems to enhance the development of materials for teaching the target genre, but, as Flowerdew (2002) notes, their findings have yet to be transferred to EAP teaching materials.

2.5 Research on students’ use of corpus technology in L2 writing

As seen above, corpus studies in L2 writing have mostly centered on materials development and syllabus design. In other words, they are mainly concerned about how to transfer corpus-based findings to teaching materials rather than giving students opportunities to use the corpus themselves. This section reviews the studies that shift the focus to learners’ own use of the corpus.

As most of L2 writers have acquired familiarity with the computer, the use of the corpus is gaining prominence in L2 writing research. The corpus as a linguistic resource
can help broaden the students’ understanding of language and enhance their writing skills. Also, the new technology may change the practice and process of L2 writing. However, researchers have not yet looked at how the use of the corpus affects students’ L2 writing behavior and performance.

Few studies have examined the effect of the corpus approach on students’ performance, which makes corpus-informed pedagogy insubstantial (Flowerdew, 2002). Only a couple of empirical studies have been done in this area. Stevens (1991) conducted the first empirical study that tested the effect of a corpus-based approach to vocabulary learning. He was interested in comparing traditional gap-fillers with concordance-based tasks in students’ performance on recalling words. A group of first-year Arab university students were tested on one gap-filler and one concordance-based vocabulary test. The contents of both tests were drawn from the students’ reading textbook. While the gap filler exercises had ten words blanked out, concordance-based exercises had ten sets of contexts for the same ten words. The results showed that the students generally showed better performance on concordance-based exercises.

Cobb (1997) replicated Stevens’ study in an online situation where students were able to manipulate a corpus-based computer program to learn new words, instead of recalling known words. The subjects were also first-year Arabic-speaking students in the same university where Stevens studied. Cobb created a computer tutor to teach a total of 240 new words over 12 weeks. He designed two versions of the program: an experimental version with concordances and a control version with no concordances. He then had the students access both versions on alternate weeks in order to measure their learning differences between the weeks with the experimental version and the weeks with
the control version. The students were given pretest, weekly quizzes, and posttest. The results revealed that the mean score of the subjects’ vocabulary levels on the six weeks with concordances was much higher than the score on the six weeks with no concordance. Also, the students demonstrated favoritism to the corpus-based instructional materials.

These two experimental studies are remarkable for their innovative undertakings. However, this quantitative measurement limited to vocabulary learning does not show a broader scope of what corpus technology does in students’ overall L2 learning and writing. Besides, the studies did not examine a variety of learner variables that may have had an influence on the results. The effect of the corpus approach can be different depending on the students’ individual differences, such as their familiarity with concordance type of task and learning styles.

Along with the test on students’ performance, it is also important to examine the students’ attitudes toward corpus technology. Ferris and Hedgcock (1998) state that “the evidence is strong that computer use improves student attitudes, confidence, and motivation and that these benefits may be even more significant for L2 writers” (p.281). Some research reports students’ positive attitudes toward corpus use in vocabulary learning, as supplementary findings (Cobb, 1997; Thurston & Candlin, 1998). In order to understand the aspects of corpus use that facilitate or impede L2 writing, we need to examine in greater depth the students’ responses to its use and any changes in their attitudes toward L2 writing as a result of their use.

Sun (2000) conducted a study that focused on students’ attitudes. He examined EFL college students’ reactions to a three-week corpus-based lesson. The lesson was
especially designed to teach the students about the corpus technique. Through a survey, Sun found that they were positive toward corpus-based learning, while some of them showed a concern about cut-off sentences in concordances and the limited corpus bank. This study presented a general understanding of the students’ responses to the use of the corpus, but it lacked a deeper qualitative insight into the topic by relying mainly upon one time quantitative data. In addition, because the study was based on a course especially designed for teaching the corpus technique, it failed to suggest how to incorporate the approach in a normal L2 writing classroom.

Turnbull and Burston (1998) conducted a longitudinal case study to examine the effectiveness of independent corpus uses by two ESL graduate students. The students used a corpus compiled from their own writings, which were corrected by the researchers. Turnbull and Burston used interviews and conference sessions, and also asked the students to write summative evaluations on their attitudes toward the use of the technique. The researchers found that unlike popular belief, the corpus approach did not always promote inductive learning strategy. Rather, individual differences, such as prior learning experience and learning styles, mediated the impact of the approach on the students’ learning. They concluded that learners need training to lead to independent corpus explorations. Besides, they suggested a further study about “the relationship between the use of concordancing strategies and language learning outcomes, and the relationship between varying degrees of concordance strategy training and learning outcomes” (p.14). The study addressed the importance of learners’ individual differences in using the corpus. Also commendable is their use of qualitative approaches that provided a deeper insight into the topic. Nonetheless, as the researchers admitted, the
study has limitations in terms of the number of participants and corpus use. In addition, the study was conducted individually, not in a normal classroom setting, thus limiting its application to use in regular curriculum.

Unlike most corpus studies that used monolingual corpora, Fan and Xu (2002) used a bilingual corpus for their research. The purpose of the study was to evaluate the effectiveness of an online bilingual corpus (Chinese and English) in the teaching of legal English in a university of Hong Kong. In particular, the researchers were interested in looking at the usefulness of the use of the corpus in the students’ self-learning. The students finished two comprehension tests on legal English while working with the online bilingual corpus on a computer. A questionnaire and a group interview were then administered in order to examine their uses and perceptions of the corpus. The results showed that the bilingual corpus promoted the comprehension of the English text. The students responded that corpus use was helpful for enhancing their understanding of the English text by giving them an option to use the two languages. While the study explored the usefulness of bilingual corpora in language education, it has limitations: 1) as the researchers admitted, it did not explain why and when the students switched between the two languages, which is essential information to be examined regarding the use of bilingual corpora, and 2) more importantly in relation to this study, the students’ corpus experience was limited to only one time exposure for an hour and 40 minutes. Thus, it does not provide a fuller explanation of the impact of corpus use on students’ L2 learning over an extended exposure and experience.

While most studies have focused on English, there has also been an investigation to deal with other languages, which contributes to expanding the scope of corpus-based
research. Kennedy and Miceli (2001) were also interested in how students become independent investigators of corpora through training. They compiled their own small corpus to teach Italian as a foreign language and led the subjects to use it as a reference in their writing. The students were also encouraged to use problem-solving approaches to revise their own writing. As a result of the study, the researchers were more concerned about revising their “apprenticeship” approach to lead the students to the independent uses of the corpus, but they also recognized the importance of formal training in the approach. This study provided valuable insight into using the corpus approach in a regular classroom. However, it used a small corpus and concordancing program that was not accessible for outsider teachers.

In summary, these previous studies have increased our understanding of corpus use in L2 writing, but they did not provide an extensive treatment of the whole issue regarding the corpus approach to L2 writing pedagogy.

First, most of the studies focused on the students’ independent uses of corpora rather than incorporating the approach as part of classroom experiences. Concordancing has been recommended as a promising pedagogical tool with which “learners explore the language for themselves and the role of instruction is to provide tools and resources for doing so” (Cobb, 1997, p.301). However, as Tribble (1991) indicates, concordancing has been used more as research strategy than as part of teaching. In other words, while much of the literature has argued for using corpora in language teaching, corpus use has rarely been observed in an actual classroom. The lack of practice is partly because corpus use has usually been initiated by researchers rather than by classroom teachers who are involved in actual teaching. In order to use the approach successfully in a real classroom,
we need to explore the questions “how can corpus-based activities best be integrated with ‘normal’ language teaching, at different levels of proficiency? How can learners (and teachers) best be trained to profit from these resources?” (Gavioli & Aston, 2001, p.245).

Secondly, few studies have explored how the use of corpora affects students’ L2 writing behavior and process. As Phinney (1996) points out, technology may not automatically generate better written products, but it may change “the way writers approach the writing process” (p.139). She adds that the focus in research on computer-assisted writing has shifted from “the question of whether using a computer helps students produce better texts” to “the changing writing behaviors and pedagogies engendered by the electronic medium, and the ways that computer technology affects how we think about the writing process” (p.139). Much needs to be done to find out how the use of corpora affects students’ L2 writing experiences as a whole.

Thirdly, little research has looked at the students’ individual experiences in the analysis of corpus use. Most of the studies assumed learners to be a homogeneous group of people with no individualistic characteristics. They applied the same approach to all the learners without recognizing different uses by individuals with different personal backgrounds. However, given the individual and private process of writing, it is also important to study writing as an individual activity of each learner. We need to develop learner-specific descriptions of corpus use in order to adjust teaching focus to each learner in the classroom.

Lastly, most of the studies used in-house programs or specialized corpora as opposed to general corpora. It is true that many scholars have emphasized the usefulness of small corpora in language teaching. However, large corpora can also play an
important role in L2 writing pedagogy. Many teachers may not be able to build up their own corpora for several reasons, such as time and insufficient skill in corpus techniques. In fact, one important reason for the lack of classroom concordancing is accessibility to a corpus and concordancing program. Stevens (1995) also indicates that the lack of the familiarity with the new technique is one major reason why concordancing is not more widely accepted in language pedagogy. It is highly likely that those teachers who are not versed in computer and corpus investigation may be discouraged from exploiting the research. Some general corpora allow free access so that teachers do not need to build up their own specialized corpora. Also, general corpora can be useful for EAP writing programs. Considering that students are often from a variety of disciplines, it may be infeasible to focus on one discipline-specific corpus in the courses. General corpora can be used more effectively by focusing the most frequent general words, thus catering to the needs of all the students in the program.

In conclusion, determining the pedagogical potential of the corpus use in L2 writing requires us to understand what it does and what changes occur in the students’ whole writing experiences. As Ferris and Hedgcock (1998) indicate, “there is much that we do not yet know about the effects of computer use on student writers” (p.267). In particular, we need an empirical report from actual teaching that uses easily accessible corpora to encourage teachers and students to go about using the new corpus approach in their own settings.
CHAPTER 3

METHODOLOGY

This chapter describes the study’s research design and procedural details such as data collection and analysis. The first section describes my initial field experience, which contributed to shaping the design for the present study. The following sections describe the overall research design, context of the study, participants, data collection, data analysis, and trustworthiness and ethics of the study.

3.1 Initial field experience

Preliminary research was conducted at a large American research university in the winter quarter of 2003. The overall purpose of the study was twofold: 1) to gain a preliminary understanding of what corpus use means to college ESL writers and how it might help them become better writers, and 2) to develop research skills and enhance the design for the present study. Given the primary interest in using corpus technology in writing, I looked for an ESL writing teacher who had incorporated a corpus approach into his/her instruction. One teacher was identified who had used corpus work extensively in his own teaching. During that time, he taught two different levels of ESL writing courses (beginning and advanced level). An initial study was conducted in both classes.
After obtaining the teacher’s permission, the two courses were observed throughout the quarter for ten weeks. A survey was administered to all the students (n=22) in both classes in the eighth week of the courses. The questionnaire asked about the students’ perceptions of the strengths and weaknesses of corpus use, and their attitudes toward corpus use in L2 writing instruction. The survey results showed, overall, that the students had positive attitudes toward using corpus technology in L2 writing. They responded that corpus use was especially useful to them for learning word usage patterns and improving their writing skills, although some of them reported that analysis of corpus data was time-consuming. In addition to the survey, follow-up interviews were conducted with nine students during the last two weeks of the quarter. The interview results indicated that individual differences, such as length of time spent learning English and the self-reported degrees of interest in L2 writing improvement, had an influence on the students’ use and perceptions of the corpus.

The preliminary study was crucial in designing the current study. First, the field experience gave me an entrée into the writing program, and secondly, it provided basic insights into the topic and communities that I sought to understand. It taught me what to expect when I, the researcher, entered the classroom for the main study. Also important is that the preliminary study gave me practical, “hands-on” experience in observation and interview techniques. Furthermore, it suggested promising avenues of inquiry. For example, it was evident that more qualitative approaches were needed in order to have a fuller understanding of the role of corpus technology in L2 academic writing. Although the study also included interviews, a one-time interview approach with each student was not sufficient to provide a deeper understanding of the topic. Besides, the study revealed
that the use of corpus technology varied from student to student, which required using methods that would focus more on the relationship between individual differences and corpus use. In short, the initial field experience helped me refine the research design to address important issues for the present study.

3.2 Design of the study

The primary purpose of this study was to examine the ways corpus technology affects ESL students’ L2 writing experience and process and to offer pedagogical insights into integrating corpora with L2 writing instruction. This study employed mainly qualitative methods, while partly integrating quantitative methods into the design. Quantitative methods alone cannot explain the whole picture of writing acquisition. The writing process is both complex and difficult to measure, and so we cannot grasp the process with only quantitative techniques. Further, an extended period of time is required to explore any change in the writing process due to exposure to new technology. Therefore, the study of writing should also include qualitative approaches, such as observations and interviews, in order to closely examine the students’ L2 writing process and learning.

The research site was an ESL writing classroom at the same institution where the preliminary study took place. Fortunately, the teacher of the study agreed to participate again, and his students participated in the study as well. It was expected that working with the same teacher would provide an advantage. Prior contact in the earlier study established significant rapport between the instructor and the researcher. The cooperation of the teacher with the researcher was a critical issue for the participant-observation
methodology used for this study. The students enrolled in the course were introduced to corpus technology and taught to use it in their L2 writing throughout the course. During this time, data were collected from all students in the class, while six case study participants were identified and became the main focus of the research. The study followed up on the six focal students in order to examine their corpus use and L2 writing experiences after the writing course.

This study used triangulation of multiple methods and data sources as a way of ensuring credibility of the data as well as obtaining thick contextualized descriptions about the topic. Three major methods employed were participant observation, case study, and survey. The methods produced six main data sources: 1) classroom observations and field notes, 2) an initial recruiting survey and a student attitudes survey from the class, 3) interviews with six case study participants, 4) recall protocols and think-aloud protocols of the six students, 5) corpus search logs and written reflections on corpus use, and 6) learning style inventories. More detailed descriptions of the methods will be presented separately in the following sections.

3.3 Context of the study

3.3.1 Site

The research site was a natural classroom setting at the aforementioned university, so as to explore the usual behaviors of teaching and learning in ESL writing courses. This university requires non-native English speakers to take an ESL writing placement test upon their arrival. The results are used to assign students to one of three courses in
the undergraduate or graduate sequences in the program. The final course in the graduate sequence was chosen for this study, as the aforementioned teacher taught the class.

The course was also an ideal choice for the purpose of this study in that, first, the teacher incorporated the corpus approach into the curriculum as part of the regular classroom activities, rather than focusing on teaching the approach *per se*, and secondly, the course emphasized the concept of writing as a process, which corresponded with a primary aspect of the study. The students in the course were asked to submit multiple drafts for each writing assignment. However, this process-guided approach of the course should not be understood as matching the traditional approach to writing process instruction because the course did not include the core elements emphasized in that tradition, such as brainstorming, peer reviewing, free writing, journal writing, and reflecting on internal resources rather than using external resources. The decision to use these writing activities was primarily left to the student writers in the research site; however, the process-guided approach the teacher adopted still helped the students gain the idea that writing involves multiple stages rather than looking at the final product as an isolated outcome.

The research site of the study can be seen as an English for Academic Purposes (EAP) writing course, rather than a general ESL course, given its content and emphasis on disciplinary writing. The course not only taught the students about the general structure of academic papers, but also required them to follow the writing conventions of their own fields. As such, students chose the topic and content of their writing based on their interests and needs in their studies.
3.3.2 Corpus

The classroom teacher in this study used a free online corpus, Collins COBUILD Corpus, which is one of the largest general corpora available. The corpus is a part of the COBUILD project led by John Sinclair at Birmingham University. The COBUILD project is in fact the largest research project by neo-Firthian corpus linguistics (McEnery & Wilson, 2001). Notably, John Sinclair’s COBUILD team is engaged in the construction of a different kind of corpus – monitor corpus – which refers to an open-ended corpus that continues to expand in size. The Collins COBUILD corpus, also known as the Bank of English, contained more than 450 million words as of January 2002. Although it is based in England, the corpus also includes texts featuring American English, and one can direct a search to only American or British texts.

Along with its free access to users, the Collins COBUILD corpus has another advantage, which is ease of use. As Rézeau (2001) notes, the corpus can be used even by novice writers, and it is appropriate to “familiarize students with concordances, as it does not require the user to be conversant with search operations” (p.150). The Collins COBUILD website [http://www.collins.co.uk/Corpus/CorpusSearch.aspx] provides a concordance and a collocate sampler from which one can draw 40 randomly chosen concordance lines and see what are statistically the most frequent 100 collocates. The sampler offers instructions on how to conduct a search, though the concordance and collocate search process requires minimal technical skill, so that anyone can learn the technique with a little practice. Also, the corpus is word-class tagged so that one can narrow a search by using the part-of-speech tags (e.g., search “use/NOUN”).
3.4 Participants

3.4.1 Student writers

A total of 14 students were enrolled in the writing course used in the study. Nine of the students were in science-related fields, two in education, two in arts, and one in humanities. Regarding nationality, the class was not very diverse; there were ten Koreans, three Chinese, and one Romanian.

As this study was mainly conducted in a qualitative framework, it sought purposeful sampling to select case study participants among the students in the class. The course instructor briefly introduced me to the class in the first week of the course. At the beginning of the second week, I explained to the class the purpose of this study and what they were expected to do if they would participate in the study. Then, they completed a brief initial survey about their backgrounds, as well as an indication of their willingness to participate in the study (Appendix A). The survey asked about their gender, major, pursuing degree, nationality, length of stay in US, length of time learning English, enjoyment of writing in English, the number of papers to be completed for the quarter, feelings about computer use, Internet accessibility at home, and previous experience in corpus use. Luckily, all of the students except two were willing to participate in the study. The information obtained through the survey was used for screening case study participants.

As one of the objectives of this study was to develop learner profiles in corpus use, participants with maximum variation were sought in an attempt to obtain a wide and rich description of the topic under investigation. The students’ backgrounds (e.g., major,
pursuing degree, gender, and age) and previous writing/technology experiences were carefully reviewed to form a diverse group of participants. Also, my presence in the class for the prior two weeks helped me to choose “information-rich cases” (Patton, 1990, p.169). Five students were selected by the second week of the study. One more student joined the study after the third week, as the teacher identified him as an enthusiastic corpus user through class assignments and tutorials and recommended him to me. Eventually, then, six students were finally chosen for the case study portion of the study.

Four of the six participants were Korean, while the remainder were a Chinese and a Romanian student. This ethnic makeup is explained by the fact that 71% of the class of 14 were Korean (10 Korean, 3 Chinese and 1 Romanian). In fact, Koreans are a sizable group of international students in ESL programs at the research site. Hence, it is not uncommon to find a Korean majority in any of its ESL classes. Secondly, the ethnic makeup is appropriate for the stated research goal of understanding the role of L1 in L2 writing. That is, it was appropriate to include a sufficient number of members of one language group for comparing corpus use behaviors among the students. In this regard, the fact that I am Korean with native fluency in the language allowed me to add personal insights based on my familiarity with the linguistic characteristics of the language. In sum, this study sought a group diverse in terms of background and experience, but also somewhat homogeneous in terms of native language so I could observe any common themes/differences among the L1 sharing students. Detailed information about the six students will be presented in the next chapter.
3.4.2 Instructor

The course instructor was an experienced American writing teacher who had taught ESL composition courses for more than ten years. Most important to this study was his interest in and application of corpus linguistics in his teaching during the two years preceding the study. In fact, he initiated use of the corpus technology in this university’s ESL composition program. As a linguist with a Ph.D. in Chinese Linguistics, the teacher was intrigued by corpus linguists’ empirical-based approach to the descriptions of language. However, his initial motivation to use corpus technology came from his experience as a classroom teacher for advanced ESL writing rather than from theoretical concerns. According to him:

I see that these students have all good GRE scores and adequate TOEFL scores. So the problems they're having are not amenable to rules. It’s not so much syntactic. It's primarily usage, collocations. Rules have gone as far as they are going to take. Really what they need is usage. ... Then, of course, at the same time, technology also became available. (Interview, 11/25/03)

The teacher’s view of language was well represented in his remark that “the primary stuff of language is conventionalized chunks, routines, and collocates.” He believed that knowing common usage patterns of words and phrases is of particular significance to advanced ESL learners in developing good writing skills. According to him, good writing is that which matches readers’ expectations about the organization and language in a given genre and ensures quick and easy information processing. He expected his students to “learn the conventions of that culture” in terms of writing gains from the course.
The teacher did not embark on using corpus technology in his teaching all at once. As he investigated the search technology available to students, he asked to what degree it would be useful for his students. He gradually increased the portion of its use in his teaching and adjusted his approach while observing students’ responses. He said that the more he used it, the more convinced he became about the utility of this form of technology in L2 writing instruction. The first and most important change he noticed in students’ writing after their involvement in corpus use was “awareness of collocations.” He observed that the students became aware of collocation and usage and started to explore those collocates themselves. For him, “it is already significant progress to have that explicit awareness, to let them question, to let them say it doesn't sound right and look it up.”

The Collins Cobuild corpus was a good choice for the teacher because of its accessibility and size. Most of all, the data in the corpus comes from various sources such as spoken, written, informal, journalistic, British, and American English. With this sizable corpus available to the students, the teacher was mainly concerned about general word usage rather than genre, because 1) “in a perfect world, we would build on genre. We create our own corpus that will be all journal articles. But we don't have that,” and more importantly, 2) “the students bring in the genre with them in the writing task.”

On the other hand, the teacher had some misgivings about students’ corpus use in terms of efficiency in analyzing corpus data, which made him adjust his focus in presenting the corpus to the students:

When I ask them to build a prototype string, I have seen them get it wrong. Or if they start to get it, they make some odd choice about a prototype string. I think it's
because they've gone too far. They have gone farther with the data than they need to go. They may have a sense they have to account for every single item on the significance list. That may be a problem. But, now, I am accumulating experience on this. Lately, with the prototype strings, I started saying, “don't account for all the data. Just do the things related to your writing task.” I think they are going to get it right with that context with the particular meaning they have in mind. (Interview, 11/25/03)

While the teacher required the students to send him weekly reports of search results of their choice, his main goal extended beyond the controlled searches for course puropses. His ultimate objective was to persuade them to take advantage of the technology as a resource on their own and to become more independent and advanced writers:

Learning, in some sense, if you accelerate learning, it seems to entail explicit awareness, set problem and then finding solution. Language is not a rule-based formal system. It's usage. So, what I hope they get is awareness of usage and appreciation of conventional language. I really try to emphasize this business of conventionalized chunks and routines that belong to certain rhetorical rules. I hope they get that too. So, that's the minimum thing, that kind of awareness. And the best thing is to appreciate this as a resource. The ultimate thing, which is the hardest thing, is link and identify usage problems. Identify usage problems on their own and solve their own problems this way. (Interview, 11/25/03)

This dissertation does not document the teacher’s views on classroom applications of corpus technology in greater detail. However, the data from an interview with him and classroom observations provided significant insights into the study in terms of situating corpus technology in a real classroom, linking the teacher’s perspectives on corpus use with the students’ expectations from its use, and particularly, interpreting the students’ corpus use behaviors and perspectives.
3.5 Data collection

3.5.1 Participant observation

The present study used participant observation methodology, as it enabled the researcher to observe behaviors and interactions in a natural setting. This had two consequences. First, it minimized the researcher’s predisposition to impose a predetermined interpretation on the data. In other words, it is an empirical approach that gave the researcher a “fresh start” on the data. Secondly, it preserved the authenticity of participant behaviors. This contrasts with a laboratory setting, which is artificial and might not produce actual behaviors that occur during teaching. The advantage of this method is that it increases understanding of the teaching and learning process. The research methodology involves observing daily classroom events in the context of the total culture of the classroom environment.

In actual practice, though, participant-observation involves more “observation” than actual “participation” so as not to disturb the teaching process (Peoples & Bailey 2003, p.105). In other words, the main purpose of participant-observation is to gain firsthand experience with an actual situation, to view behaviors as they occur, rather than personally involving oneself in the dynamics of that situation. I did have an advantage, though, as I was a former student in a writing class in this program, so I also brought an insider quality to this study. This insider knowledge provided me with certain emic perspectives on the inquiry. But I observed and participated in every class for ten weeks so as to become a more “genuine” insider of the specific classroom culture investigated in the study. At the same time, I was an outsider, as I had been trained in L2 writing
pedagogy through my own graduate studies, which helped me wear a teacher’s lens. This “insider out” perspective gave me a wide range of understanding of the phenomenon being studied. In particular, participant observation methodology helped me reflect on my experiences in the writing program, and also allowed me to critique my own training in L2 pedagogy.

Angrosino and Mays de Perez (2000) state that the most important contemporary use of the observer-as-participant role is in classroom observational studies. As they indicate, I expected my role to shift from “complete observer” to “participant” on a continuum from time to time. Although the participation was limited, it still reduced the social distinctions between the researcher and the researched and helped me establish an insider relationship with the participants. It allowed me to build rapport and a sense of camaraderie with them. Rapport was necessary in order to obtain students’ candid appraisal about the corpus approach.

During observations over the ten weeks of the quarter, personal reflections as well as observational notes were kept in a researcher journal. Classroom observations not only described the instructor’s teaching practice, but also yielded a significant portrait of students’ experiences in the classroom. At the same time, the method generated a way of looking at students’ use of corpus technology within the whole context of writing instruction. Through the firsthand observation of the classroom, I was particularly interested in exploring what activities were constructed for corpus use, the interactions between teacher and student and the resulting dynamics, and how those dynamics fostered learning over time. Importantly, the participant observations also helped identify
participants for the case studies and enhanced development of the rapport necessary for the case studies through regular contacts over the course of the research.

3.5.2 Case study

The major methodology used in this study was case study, “which has proven to be the most effective way to examine the writing process” (Zamel, 1983, p.169). It views research as a meaning-making process in which researchers understand and interpret the multiple perspectives that people bring to the phenomenon in question. The methodology does this by capturing the experience as lived by the participants (Janesick, 2000). It is necessary then to pursue long-term interaction with the participants to obtain in-depth understanding, and thus it is critical that the researcher establish rapport with participants.

The present study conducted multiple case studies with six students whose data provided students’ voices to the descriptions of the role of corpus technology in L2 writing. Each case study included interviews, recall protocols, think-aloud protocols, and document analysis. These methods were designed to produce thick descriptions of their experiences in corpus use and its influence on their writing.

Once the participants for case study were identified, they were first asked about their writing processes by use of interviews and recall protocols. The data were used as a baseline to determine any changes in their writing processes after corpus use. Also, their learning styles were examined at the beginning of the study, as learning styles are believed to play an important role in adapting to the corpus approach (Turnbull & Burston, 1998). Instruments that assess learning styles were employed to complement
the students’ self-reports on their learning styles. Two inventories were used: Style Analysis Survey (SAS) and Strategy Inventory for Language Learning (SILL). Both inventories were intended to complement each other, as SAS examines overall learning styles regardless of subject and language, while SILL explores language learning strategies for non-native English speakers.

The SAS is a learning and working style survey designed by Rebecca Oxford in 1989. It purports to identify overall style preferences. It has 5 sections. Students are required to mark their responses to statements in each section, using a scale of 0 to 3 (0 = never, 1 = sometimes, 2 = very often, and 3 = always). Then the scores are tallied. The highest score is considered their style preferences. If there are scores within 2 points apart, both are considered their styles. On the other hand, the SILL examines the strategies that students use in their language learning. The SILL used for this study was the version for speakers of other languages learning English also designed by Rebecca Oxford in 1989. The inventory contains 6 sections and the students are asked to respond to each statement of the sections by using a scale of 1 to 5 (1 = never or almost never true of me, 2 = usually not true of me, 3 = somewhat true of me, 4 = usually true of me, and 5 = always or almost always true of me). At the completion of the inventory, the scores are added up and divided by the number of statements in each section. The average scores show the degree of use of a corresponding strategy. That is, an average of 4.5 to 5.0 = always or almost always used, 3.5 to 4.4 = usually used, 2.5 to 3.4 = sometimes used, 1.5 to 2.4 = generally not used, and 1.0 to 1.4 = never or almost never used. A sample SAS and SILL are enclosed in Appendix B and C, respectively.
The participants were interviewed approximately once every two weeks for an hour during the ten weeks of the course. Due to a corpus service breakdown that occurred before the start of the following quarter, only a couple of interviews were conducted after the course ended. In terms of format, semi-structured interviewing was used; while common thematic questions were asked across the cases, room was left to improvise in individual cases and to respond to the uniqueness of each case (Fontana & Frey, 2000). Throughout the research period, the participants were mainly asked about their experiences in using the corpus: their actual use of the corpus, their evaluations of its use, and any changes in their writing and writing process after corpus use. The students’ self-reflections on their writing process associated with corpus use yielded introspective data about the influence of corpus technology on their understanding of language and their approaches to L2 writing that otherwise would not be accessible. The interviews were recorded on audiotape and transcribed as soon as possible in a standard word-processing program for subsequent analysis. In addition to the formal interviews, informal spot interviews were conducted after class when a significant issue regarding corpus practice emerged unexpectedly in a class. Also, frequent informal conversations with the students outside the class not only built rapport, but also helped me to better understand their life experiences and contexts of learning.

Think-aloud protocols were employed in order to examine the participants’ processes of corpus data analysis. They were given a short practice session before the actual task to become familiar with the task. A set of concordance and collocate data was then given to them to differentiate usage between two similar words (“use” and “usage”).
The think-aloud task was intended to observe the students’ actual corpus use patterns and strategies in analyzing corpus data.

This study also collected a variety of documents (i.e., the participants’ corpus search logs, class e-mail assignments, and written reflections on corpus use). Those data revealed their frequency of corpus use, purpose of corpus searches, and common search items. In particular, the students kept a log of their corpus searches, which they brought to the following interview. In the log, they recorded search items, the purpose of searches, their level of satisfaction with the search results, and the amount of time spent on the searches. The log form is attached in Appendix D.

### 3.5.3 Survey

A survey was administered in the last week of the course to all students in the class in order to examine the class’ overall amount of corpus use and attitudes toward its use. Out of a total of 14 students, 13 participated in the survey, one being absent. The original questionnaire was constructed during the preliminary study based on an examination of the literature (e.g., Sun, 2000) and in consultation with the classroom teacher as well as an L2 writing scholar, who has expertise in the teaching and research of L2 writing. The survey instrument was field tested for suitability and content validity in the preliminary study. As the present study proceeded, a more contextualized questionnaire was created to reflect the emerging themes and patterns captured through data collection and analysis. In this way, it was a kind of grounded survey based on the insights obtained from empirical work in the research context. This is a feature of qualitative research. The survey questionnaire is enclosed in Appendix E.
The questionnaire consisted of seven main parts: 1) computer and dictionary use, 2) perceptions of English writing, 3) advantages of corpus use, 4) difficulties/problems in corpus use, 5) corpus use in English writing, 6) corpus use in writing instruction, and 7) demographic information. Except for information about computer and dictionary use and demographic information, the questionnaire employed Likert scales that were composed of statements participants responded to. The students were asked to use a scale to indicate their level of agreement with the statements. The questions involved simple and direct language to represent clear ideas to the respondents.

3.6 Data analysis

Regarding data interpretation/presentation, the issue of translation should be first considered. This study included four Korean case study participants who shared the same native language with me, which produced bilingual data. The participants were interviewed in Korean by their choice. This study benefited from the bilingual research process by providing rich data and promoting a better understanding of the students’ perspectives. On the other hand, the bilingual aspect raised a question regarding reliability of the translated data. The Korean interview transcripts needed to be translated into English, with the caveat that the participants’ voices were interpreted and translated into words of the researcher’s choice, rather than presenting their verbatim accounts.

The issue of representation of others is a critical point even in transcribing and interpreting interviews, let alone translating them into another language. As Hodder (2000) put it, “once words are transformed into a written text, the gap between the ‘author’ and the ‘reader’ widens” (p.704). As such, an inevitable question arose as to
whether my translation accurately represents the participants’ perspectives. Still, careful attempts were made in order to do justice to the participants’ words. The interview data from the Korean participants presented in the following chapters are my translations from Korean into English, unless they are marked as “original in English.” Excerpts of translation are enclosed in Appendix H.

As is common in qualitative research, the data analysis was intertwined with the process of collecting the data in this study. Analysis of the data components was done simultaneously with data collection so that the study was shaped to focus on issues emerging as data were collected. Qualitative research is an act of interpretation from beginning to end (Janesick, 2000). While most of the data was largely textual, data reduction for each data set generally involved the following steps: 1) skimming through the data, 2) second reading to identify highlights that illuminate the broader context, 3) third reading while coding, 4) finding emergent themes and patterns across the coding, and 5) looking at the data sets again within the identified themes and patterns.

While observational notes were used as baseline data for the analysis of the whole picture of the classroom and of teaching practice, the case study data sets were analyzed in order to identify commonalities and differences among the participants. The analysis of interview data was focused on identifying their perspectives on L2 writing, writing approaches and processes, and individual experiences and learning contexts that affected their corpus use and writing experiences. Once the participants completed the SAS and SILL inventories, the results were produced by the inventories’ instruction and used to examine the relationship between the students’ learning styles and their corpus use patterns. The results were also provided to the students. Corpus search logs were
analyzed to identify the frequency of their corpus use and their search items, and the purpose and results of their searches. The analysis provided important information about the linguistic features in which the students had difficulty in their writing and about their approaches to corpus use, such as focus on form or on context or both. Summary tables were prepared and analyzed.

The class survey that produced quantitative data was analyzed by use of statistical tests (SPSS Version 11.0 for Windows). The Likert type responses were summated and treated as interval data. Descriptive statistics produced results such as frequency, mean, and standard deviation. While this dissertation mainly presents the findings of the case studies, the results of the survey are provided in Appendix I.

3.7 Trustworthiness and ethics of the study

Prior to the beginning of this study, a human subjects review proposal was approved by the Behavioral and Social Sciences Investigative Review Board (IRB) at the research institution. Under IRB guidelines, I obtained informed consent from the six case study participants, while also obtaining permission from the class as well as the teacher for classroom observations. The consent form and recruitment letter are enclosed in Appendix F and G, respectively.

Apart from IRB procedural approval, I also attempted to enhance this study’s ethical performance by creating a reciprocal relationship with the participants. Researchers often suddenly withdraw from the field site and subjects once their data collection is completed. From an ethical standpoint, researchers must ask what benefits accrue to the participants for their time and effort in the study. The lack of reciprocity
can result in the appropriation of the participants by the researcher. I developed a close relationship with the participants throughout the study, and we maintained the relationship (actually, friendship) even after the study. They often asked me, an old-timer in the environment, for information and sources regarding academic and everyday lives in America. Besides, I helped them by proofreading their term papers. Most of all, they appreciated the opportunities to reflect on their L2 writing and discuss their difficulties and possible ways to improve.

The ethics of the study are closely related to the truthfulness of the findings in a qualitative design. Qualitative research uses the notion of “trustworthiness” to judge the quality of inquiries, which follows different evaluative criteria from those used for quantitative-based methods. Quantitative inquiries use four evaluative criteria – internal validity, external validity, reliability, and objectivity – which are best ensured by controlling environments and warranting freedom from the contamination of the researcher (Donmoyer, 1990). The qualitative paradigm questions the traditional basic assumption of researcher objectivity and neutrality and thus requires a serious rethinking of the evaluation of research. Guba and Lincoln (1989) developed a set of evaluative criteria parallel to the four conventional criteria: credibility, transferability, dependability, and confirmability. These were the criteria used in this study.

In order to enhance credibility, the present study used triangulation of multiple methods and data sources, prolonged engagement, and persistent observation. The use of multiple methods in qualitative research not only enables an in-depth understanding of the phenomenon, but also helps to establish trustworthiness of data. This study combined multiple data gathering methods: observations, interviews, a survey, and document
analysis. It also attempted to include multiple perspectives from the classroom teacher as well as multiple case study students with diverse backgrounds. As the researcher is the main research instrument of qualitative inquiry (Janesick, 2000), triangulation of multiple sources and perspectives served to guard against my personal biases. My prolonged interaction with the students also helped increase the study’s credibility. Qualitative research involves prolonged engagement in the field site and participants, enabling the researcher to produce a rich description and enhance the representation of the participants. Further, my persistent engagement in the topic, including the preliminary study, as well as insider experience in the program as a former student contributed to the trustworthiness of the findings.

From a traditional research paradigm, this sort of study should be conducted in a statistically significant number of randomly selected classrooms over, say, one year, to obtain a full picture of L2 students’ writing practice over time, so the limitation of this study is the ability to generalize. However, qualitative research does not make claims about its generalizability to a larger population. Instead, it attempts to provide rich and theorized descriptions of the immediate context, aiming for transferability that “helps the reader judge the applicability of the findings to new contexts” (Morita, 2002, p.74). In this study, the thick descriptions of multiple cases with a wide range of backgrounds helped enhance the transferability of its findings.

Finally, I attempted to enhance dependability and confirmability of the study through an explicit documentation of method decisions, member checks and a reciprocal relationship with the participants. In particular, member checking is an important method to increase trustworthiness of the data in a qualitative inquiry. The participants were
provided with opportunities to check my interpretations of their beliefs and behaviors through our ongoing conversations and interviews. In this way, member checking was used not only as a way of checking assertions made in this study, but also as a form of reciprocity with them. In qualitative research, the research process is a negotiated interaction between researcher and researched. Lincoln (1995) argues that “detachment and author objectivity become barriers to quality, not insurance of having achieved it” (p.280). While quantitative inquiries view the researcher as a potential contaminant to be separated out and controlled, qualitative researchers do not discover objective reality; rather, they construct social reality and make meanings through interactions with the researched (Denzin, 2000; Fine, 2000; Lather, 1986). In this study, I established and valued intimate relationships between researcher and researched. I attempted to use the data to assist the participants in their learning and also discussed the findings with them, which offered to me as well as to them opportunities to reflect on L2 writing. In this way, we were collectively engaged in constructing new knowledge, and the reciprocity contributed to enhancing the trustworthiness of the study.
This chapter portrays significant features of the six case study participants’ background and writing experiences. It is an important backdrop for the next findings chapter, as this is critical to providing the context for the ethnographic/qualitative aspect of the study. The first section, 4.1, describes the focal students’ demographic data and personal histories. While Section 4.1 introduces the students case by case, the following sections present their views on relevant topics through across-case analyses in order to provide their collective perceptions and patterns as well as identify individual voices. Those sections include: 4.2 views on language and language learning, 4.3 their approaches to L2 writing, 4.4 their difficulties in L2 writing, 4.5 their L2 writing processes, and 4.6 their learning styles, which were all explored at the beginning of the study.

4.1 Introducing case study participants

Out of the class of 14 students, a careful attempt was made to choose a diverse group of subjects for the in-depth qualitative analysis for this study. The selected group of six was chosen to reflect diversity in terms of age, gender, academic major, writing
experiences, and technology skills. These students’ varied educational and contextual backgrounds become apparent in the in-depth analysis of their responses to the corpus, and to their perspectives on language learning and writing. In this way, we have confirmed the value of integrating the qualitative approach with quantitative methodologies. In other words, this is not a matter of “either – or” in evaluating methods, as is often the case in discussions of mixed modes studies. Rather, we demonstrate that integrating both methods provides substantial insights that could only be partially attained by using just one method (i.e., qualitative or quantitative).

This is an appropriate point, then, to introduce the six case study participants by describing their personal histories. Research subjects in any study bring a reservoir of complex backgrounds not easily reducible to simple variables. These complexities influence their responses to interview questions. By understanding each of their personal histories and backgrounds and probing the interplay of personal attributes and learning contexts, we can better understand the students’ reactions to the questions.

Table 4.1, “Overview of case study participants,” provides this important overview of critical life experiences and learning contexts which were considered to be most likely relevant to this study.¹

¹ The names for the participants are pseudonyms to ensure confidentiality.
<table>
<thead>
<tr>
<th>Name (age, gender)</th>
<th>Wol (28, F)</th>
<th>Chan (29, M)</th>
<th>June (24, F)</th>
<th>Sung (32, M)</th>
<th>Nick (29, M)</th>
<th>Xiaodong (29, M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native language</td>
<td>Korean</td>
<td>Korean</td>
<td>Korean</td>
<td>Korean</td>
<td>Romanian</td>
<td>Chinese</td>
</tr>
<tr>
<td>Program of study</td>
<td>MEd, Special Education</td>
<td>Ph.D., Natural Resources Education</td>
<td>MS, Aerospace Engineering</td>
<td>Ph.D., History</td>
<td>MS, Nuclear Engineering</td>
<td>Ph.D., Molecular Genetics</td>
</tr>
<tr>
<td>Years of English learning</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
<td>5</td>
<td>10+</td>
</tr>
<tr>
<td>Months in U.S. at the start of the study</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>TOEFL (TWE) scores3</td>
<td>240 (4.0)</td>
<td>270 (4.5)</td>
<td>240 (3.5)</td>
<td>253 (4.0)</td>
<td>217 (3.5)</td>
<td>607 (3.5)</td>
</tr>
<tr>
<td>Unique experiences affecting English/literacy skills</td>
<td>Korean language teacher (3 ½ years)</td>
<td>Military service at US Army base</td>
<td>Qualitative research thesis (MEd)</td>
<td>MA in Japan</td>
<td>Took previous writing classes in the ESL program</td>
<td>Published a paper in English</td>
</tr>
<tr>
<td>Enjoy writing?</td>
<td>Yes</td>
<td>Yes</td>
<td>No (even in L1)</td>
<td>No</td>
<td>No (even in L1)</td>
<td>No</td>
</tr>
<tr>
<td>L1 academic writing experience</td>
<td>Experienced</td>
<td>Very experienced</td>
<td>Semi-experienced</td>
<td>Semi-experienced</td>
<td>Inexperienced</td>
<td>Experienced</td>
</tr>
<tr>
<td>L2 writing experience</td>
<td>TWE</td>
<td>Took a class in Korea &amp; TWE</td>
<td>TWE</td>
<td>TWE</td>
<td>1 yr coursework in US &amp; previous writing classes</td>
<td>Published a paper &amp; TWE</td>
</tr>
<tr>
<td>L2 writing confidence?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes (very high)</td>
<td>Yes (very high)</td>
</tr>
<tr>
<td>Content knowledge</td>
<td>Low</td>
<td>High</td>
<td>Intermediate</td>
<td>Intermediate</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Technology skills</td>
<td>Low</td>
<td>Intermediate</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Like to use computers?</td>
<td>No</td>
<td>No</td>
<td>Medium</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Initial attitude to corpus</td>
<td>Positive</td>
<td>Positive</td>
<td>Unsure</td>
<td>Negative</td>
<td>Very positive</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Table 4.1: Overview of case study participants

---

3 Student approximated years of study

Xiaodong’s score is based on the former pen-and-pencil test, while the others are based on computer-based tests, the most current format of the test.
In terms of programs of study, the students represent a broad array of majors across the academic spectrum. Four students were majoring in science-related fields (one in combined science and education), one was in education, and the other one in the humanities. Two participants, Wol and Nick, had already been in the United States for one year; the other four students were new arrivals, and they were still adjusting to the new environment as foreigners. All except Nick were experiencing their first academic term in an American university setting. Wol had been in the United States for the previous year as a housewife, but this was her first American academic term.

The students were assumed to be at similar, if not the same, levels of writing proficiency because they were assigned to the same course based on placement test results. However, Nick was an exception. He had already spent one year in the institution where he took the first two courses of the three course-sequenced writing program. During that time, Nick also performed many writing assignments in classes and learned L2 writing by virtue of living in the day-to-day experience of the United States.

**Wol**

Wol was a 28-year-old Korean female master’s degree student in special education. Actually, she volunteered to participate, expressing much interest, immediately after I explained my agenda to the class. My first impression of her was that she was friendly, calm, and confident. In the initial survey, she indicated she did not like to write in English, yet she had the highest number of writing tasks to be completed that quarter in her studies. She became a valuable participant because she had been a (Korean) language arts teacher at a junior high school in Korea for 3 and half years.
Thus, this teacher-turned-learner qua research subject provided layers of pedagogical insight into this study, particularly as it examined students’ L2 writing experiences while considering their L1 experiences.

Before going to college, Wol studied English for six years in junior high and high school. However, she did not like the school’s approach because it was primarily devoted to learning grammar and reading. Instead, she was more interested in an approach that developed interactive communication because, for her, “language is not knowledge to be acquired, but a habit to be formed through using the language in a real situation.”4

Wol’s unsatisfactory English learning experiences convinced her to major in Korean language and literature in college, because she did not like English. Due to the importance of English as an international language, however, she pursued further study of it after high school. She attended language schools to learn English on her own. But, this time, the approach changed from learning a “dead language” (죽은 영어) as subject matter to learning a “live language” (산 영어) to communicate with people. Her special interest in communication was due in part to her experiences in traveling the world.

After college graduation, Wol taught Korean language arts until she came to America while her husband studied in engineering for his doctoral degree. During her first year in the United States, she was a housewife adjusting to American culture. She took a few free English classes offered to the wives of international students. Those

___________________________

4 Four Koreans, Wol, Chan, June, and Sung, were interviewed in Korean by their choice. Thus quotations from them were my translations from Korean into English. On the other hand, the other two non-Korean students, Nick and Xiaodong, were interviewed in English, and quotations from them were verbatim transcriptions.
classes were not entirely satisfactory because they focused on grammar rather than conversation. However, she enjoyed making international friends, meeting American teachers, and sharing cultural knowledge. As such, the cultural aspect was something she wanted to acquire through language.

After one year, she decided to pursue a master’s degree and chose to enter the special education program, which was a new field for her. Before coming to America, she did not have a clear idea about what she would study. But, for her, a major area did not have to be related to a future career or prospect of job opportunities. She was on leave from her school and had the security of going back to the school at any time. She believed that her new learning would help her to teach students with special needs. The master’s program required students to take an exit examination, rather than writing a thesis for partial fulfillment of graduation requirements. She did not plan to pursue a Ph.D. degree right after her master’s degree.

Having one year of acclimation to the USA, she seemed to start her study with less of a burden compared to the others, who had to adjust to the new environment. While her personal goal was improved conversation skills, she did recognize the importance of developing writing skills when she started her studies.

With regard to writing experiences, Wol enjoyed reading and writing in her native language, as she used to be a Korean language arts teacher. She also had extensive experiences in Korean academic writing. As a Korean language arts major, she had to write a great deal in college, almost an essay every week. In contrast, she did not have any academic writing experiences in English. Like most of the other participants for this study, her first and only experience in L2 writing (in the real sense of composition, not
merely the translation of texts) was through the Test of Writing in English (TWE), which is included in the Test of English as a Foreign Language (TOEFL). In addition, she neither had much content knowledge nor disciplinary writing experience in her field of study, since the major was a new field to her.

Interestingly, my field notes and interview data revealed that Wol was not experiencing great difficulty in L2 writing. On the contrary, she displayed competence in L2 writing. The reason why she was dissatisfied with her L2 writing skills was related to her high expectations as a competent writer in Korean:

I don’t dislike writing in English because I was a Korean language arts teacher and I liked to write anyway. But I don’t like the fact that I can’t express myself fully in English as much as I can in Korean. I want to keep my own unique voice in English, but I can’t. (Interview 1, 10/14/03)

Regarding technology, she did not like to use a computer very much. Reading a book was more enjoyable to her. Yet she did not deny that the computer is indispensable for her academic work and for communication with others. Also, she showed a positive attitude to corpus use. She reported that when the teacher first introduced the corpus to the class, she exclaimed “Wow! This is exactly what I want.”

**Chan**

Chan was a 29 year-old Korean male student pursuing his doctoral degree in natural resources education. He joined this study later than the others. The course instructor recommended Chan to me after observing his enthusiastic corpus use. His

---

5 Wol took the TOEFL and GRE test one year before she started her graduate studies.
recommendation came three weeks after the study was underway. Through class assignments and conferences, the teacher found that Chan was actively using the corpus and showing a great deal of interest in the new technology. With the teacher’s encouragement, the student agreed to participate in the study.

As a research participant, Chan was very energetic, inquisitive and articulate. He was often outspoken in the class. This study benefited from his participation, because he often elaborated on answers to questions, always making sure he got his point across and that it was understood. He provided more than enough information in the interviews. Throughout the study, he was an ideal informant who provided the in-depth “insider’s” view toward his personal exposure to corpus use and L1, L2 writing.

At college in Korea, Chan majored in science education with a specialty in chemistry. After graduation, he taught science at a junior high school. While teaching, he continued his studies in the same program for his master’s degree.

Like Wol, Chan enjoyed writing in general, and among the participants, he had the most experience in L1 academic literacy. He wrote many papers during his graduate studies. In addition, he wrote a qualitative research thesis in partial fulfillment of the requirements for his master’s degree. He commented:

I think I am kind of addicted to reading. I often just read any kind, a book or Internet website. Because I like to read and write, and play with words, the qualitative thesis writing was something that I greatly enjoyed. (Interview 2, 11/10/03)

Unlike the other students, Chan was already exposed to English in a natural (not classroom or artificial) situation before he came to America. He did his military service
at a U.S. Army base in Korea for two years. His military experience probably enhanced his speaking skills, but he also took an English academic writing course during his master’s program in Korea. It was, then, not surprising that he showed confidence in L2 writing. Upon entry to the program, Chan had a TOEFL score of 270 out of 300, with 4.5 in TWE. Most importantly, he stated that he liked writing in the L2. Accordingly:

I like writing especially here in America because I can’t express myself well in speaking during discussion in class. I can express myself much better in writing because I can take more time to prepare, and I can check whether it’s correct. However, I still don’t think I can fully express myself even in writing. Maybe 50% of my ideas can be expressed in writing, but only 20-30% in speaking. (Interview 1, 10/28/03)

Moreover, pursuing the same major throughout his college and graduate studies provided him with a great level of content knowledge in his field.

Notably, Chan had a great interest in linguistic analysis from semantic levels (e.g., analyzing a Korean word about its origin) to syntactical features (e.g., metacognitive awareness of language). These factors became evident when he expounded on his analysis. However, he reported that he focused more on the flow of ideas in written discourse, rather than on surface features in writing.

Chan admitted he did not like using computers. However, he used a computer frequently for his academic work and for pleasure reading. Mostly he used the computer for word processing, seeking information on the Internet, and for using electronic dictionaries. Pleasure reading was an addition to the list after he came to America in order to get information about Korea. With regard to the corpus, like Wol, he showed a positive attitude towards its use.
June

June, 24, was a Korean female majoring in aerospace engineering. The youngest participant, she had just graduated from a Korean university and had recently come to America for her master’s degree. In the class, she gave the impression that she was quiet, shy, and introverted. In the interview, however, contrary to my first impression, she was talkative and verbally animated.

June stayed with a host family for the first 10 days in the United States, giving her an opportunity to practice basic spoken skills and to get exposed to “real American culture.” She seemed to be very sensitive to the new culture. Throughout the study, she expressed interest in learning American culture, which reflected and affected her perspectives on the language and language learning. Cultural acquisition, for her, was a necessary prerequisite for becoming fluent in the given language. As she recognized the value of language learning in cultural acquisition, she preferred to learn a language in its natural environment rather than in a classroom context. She created more opportunities and found resources to gain more exposure to the language and the new culture. For example, she kept in contact with her host family on a regular basis and went to church with them on Sundays.

June started learning English in elementary school as an extra-curricular activity. But she did not remember much about the experience. Then she studied English throughout her secondary school as an academic subject, though not in the sense of an interactive communication skill. Being in the science-track in high school, she did not

---

6 Korean high school students are required to choose either humanities/language arts track or science track in their second year (corresponding to U.S. 11th grade). The choice is made in relation to their
focus on Korean and English language arts as much as on math and science subjects. For her, English was no more than an academic subject, like world history, which required her to absorb knowledge (e.g., grammar rules and vocabulary). Yet her perspectives on English learning changed in college where she encountered “meaningful” English texts related to her study. The applied content motivated her to study more:

After going to college, I started feeling the need for learning English to read textbooks in English and I realized that many people in the rest of the world learn English. Since then, I began to become more active in studying English. I thought I needed to learn their way of thinking. Then, I felt my way of thinking was changing. Learning a language can’t be separated from its culture, so studying English advances my academic and cultural knowledge of the language.

(Interview 1, 10/15/03)

In the first interview, June said she struggled with listening to and speaking English in America. Her foremost desire was to be able to speak her opinions in the content courses. Therefore, she was eager to learn English and was very self-conscious about her language learning. She even tried speaking in English with her Korean colleague in her laboratory and with her Korean roommate to enhance her spoken English skills. However, she confessed that expressing herself in writing was even more difficult. Overall, she was highly self-critical and harsh in evaluating her English language competence in general, and her writing skills in particular. She joked that her TWE score (3.5) was “the lowest score that anyone with no preparation of the test could earn,” even though Korea’s educational system does not prepare students for composition.

future possible majors in college. The two tracks share core subjects, but they also have different foci and subjects in their curriculum. So, it is unusual to change a major track in the future as their learning is directly related to the college entrance examination.
In general, June did not like writing. For June, writing was just “stressful,” as she had no successful experience in writing, even in Korean. As a science major, she was not given many opportunities to write papers. Most of her assignments were about computer programming, and she wrote reports that were preoccupied with numbers and mathematical formulae. Thus, she was not confident in writing in her L1, let alone writing in her L2. Except for her TWE experience, this class was the first time she ever engaged in academic writing in the L2. However, she had a strong desire to develop writing skills that would allow her to have a successful master’s course of study and eventually write a good thesis. She was also motivated by a desire to write a general-interest book in the future. To be a good English writer, she was convinced that she had to adopt the American ways of thinking and expression based on her firm conviction that cultural and language acquisition were interdependent. Regarding her depth of knowledge in her discipline, she was somewhat confident because she majored in the same field as an undergraduate, but she still felt her knowledge base was not “deep enough.”

June had a high level of computer skills because of the computer’s use in engineering studies. On the other hand, she demonstrated only a medium level of interest toward corpus use. She was unsure about how she could benefit from the corpus, but she was still open-minded enough to try it.

The corpus is just interesting to me. But I can’t find what I want. I am not sure why. Maybe, I still don’t know exactly how to search? But I can’t find the expressions that I want to put in my writing. (Interview 1, 10/15/03)
June did not question the value of corpus, but found it unhelpful for her particular needs or learning level. She was hampered by her limited knowledge of the searching query. An alternative explanation for her difficulty became apparent in the study, which will be reported later in this dissertation.

**Sung**

Sung, a 32-year-old male, was from Korea. He was pursuing a Ph. D. in military history. He had a diverse background in terms of life experiences, languages and academic studies. He graduated from the Korean Army Academy, where he majored in Japanese linguistics. As a professional military officer, his graduate studies were supported by the Korean Army. For his master’s degree, he went to a Japanese university to study international politics. Then, he returned to military duty in Korea until he came to America for his Ph.D. studies in history. While history was a new field for him, his focus was on military history, in which he was knowledgeable from his professional career. In his first quarter of study, he already had chosen his dissertation topic: an analysis of the Japanese army and navy in the Pacific War.

At the risk of stereotyping his personality, Sung seemed to have a military officer’s characteristics. He was friendly, yet reserved and determined. He was reticent in our interviews. In sharp contrast to Chan, Sung usually gave brief answers, mostly yes or no, which required extensive follow-up questions.

Sung studied English in secondary school similar to the way many other Korean students did: focusing on grammar and basic reading. His experience in college, however, was different. While most of the Korean universities maintained the focus on
reading and grammar, the army academy's curriculum emphasized spoken English because of the need for joint training with the US Army. Sung recalled he earned better scores in the listening section rather than for the grammar section in the TOEFL, which was the opposite of most of the Korean students. Of particular interest here was his negative attitude toward the spoken-focused English program in college. He wished he had had more grammar instruction, which was exactly the point the others believed to be the problem (or the reason for their communication shortcomings) in their English learning experiences. Sung often expressed his need for more grammar and accuracy during the study.

Like June, Sung did not like writing in general. In his college, he was not given many chances to write academic papers in Korean. Most of the assessment was through tests rather than in writing papers. Interestingly, he recalled that his foremost academic writing experience was in Japanese. In fact, he wrote his master’s thesis in Japanese. So English was the third language for him. With respect to English writing, the TWE was his only opportunity to produce a long English text that was not for translation, but for composition in a real sense. Even though he started learning English earlier and studied it longer than Japanese, he felt much more comfortable with Japanese than with English. As to why, he responded:

First of all, Korean and Japanese have the same word order. I am not saying writing in Japanese was easy. But at that time I was not so concerned about making grammatical mistakes because I believed that I had no problem making myself clear and conveying meanings. In contrast, now in the States, I am afraid that I may not make myself understood, which is actually happening, and I am not happy about getting a bad grade for it. (Interview 2, 11/4/03)
As can be seen, he was very concerned about his grades. Later I found that his grades were one of the most critical factors in his prospects for career promotion. I was informed that the Army does value grades rather than other academic achievement, such as publications. Overall, Sung was experiencing a harder time in America than in Japan. He commented that Japanese people seemed to be friendlier to foreigners than Americans, and they were more tolerant toward foreigners’ mistakes, which made it easier for him to communicate in Japanese than in English.

Sung did not like computers. He used a computer as required for personal and academic pursuits, but he did not like to spend time playing on the computer; he would rather read a book. He used a word-processing program for writing and read Internet news articles regarding Korea that were not accessible through newspapers in America. Regarding the use of corpus technology, he was the only participant who clearly showed a negative attitude at the beginning of the study. He remarked:

For now, I wonder why I need to use it. Examples in the Korean-English bilingual dictionary are good enough for me. I know the corpus may have an advantage, but I still don’t see any reason why I should change to it all. (Interview 1, 10/21/03)

As we can see, both June and Sung were not positive toward corpus use. But the different reasons for their attitudes are worth noting. While June blamed herself for not knowing enough to use the search functions, Sung questioned its advantages over the dictionary. In other words, June was not skeptical of corpus use benefits, but she needed additional help in negotiating the technology. Sung, on the other hand, remained skeptical about the approach, regardless of his ability to use the technology. These are
important distinctions worth keeping in mind. We will revisit this point later, because it will show how their initial attitudes influenced their corpus use behaviors.

**Nick**

Nick, 29, was a Romanian male majoring in nuclear engineering. At the beginning of the study, Nick had already spent one year in his master’s program. As mentioned earlier, he had also completed the first two courses of the three-course writing program. Since that program incorporated corpus use into the curriculum, Nick was the only participant who had already been exposed to the corpus. In fact, the instructor of this study taught his first ESL writing class, and at that time, Nick participated in the prior study preceding this project. His participation was critical because his case provided an extended period of corpus use. Another advantage was that we had already established rapport in the earlier study. Therefore, he was chosen as a default. When I explained the research project, Nick immediately volunteered to participate again. In the prior study, he expressed a highly positive attitude toward corpus use. He extolled the benefits of the corpus for learning the contexts of words in general and for applying prepositional usage in particular (Yoon & Hirvela, 2004).

Nick was outspoken in class and often voiced his opinions to the teacher. Nick also had a great deal of fluent spoken language at his disposal. Yet, Nick was the participant with the least amount of years of English learning before coming to America. Nick studied English for only 3 years in secondary school in Romania. The classes were mostly devoted to learning grammar rules, with minimal emphasis on listening comprehension. Then he received a college scholarship to study in Belgium, where he
did not study English. After graduation, he resumed studying English to prepare for the TOEFL and GRE in anticipation of studying in America.⁷ Thus, he had been studying English for about 5 years. Nick commented that he did not encounter much difficulty in his English learning because, he said, he found English to be similar to Romanian.

Nick did not like writing in his native language, and he did not have any experience in L1 academic writing. Regarding L2 writing, he enjoyed informal correspondence, such as by e-mail. He also had great confidence in writing in English, though he did not enjoy it. For him, writing in English was not difficult. He said his confidence came from his learning and his experiences in writing courses. He frequently reiterated how helpful the courses were in developing writing skill:

I write it because I have to write it. I don’t like writing in English. I don’t like writing anyway even in Romanian. …. Difficulty in English writing? Um… [It depends on] what you want to say, but if I have an idea in my [mind], it’s not difficult to express my idea. I don’t know. It’s not a problem. Maybe I can make some mistake, but I can fix it, so no problem! At the beginning it was a disaster. I didn’t have [any] idea about how to write, what steps I needed to follow, but now I know more after the composition classes. (Interview 1, 10/24/03)

To reiterate, Nick had the fewest number of years of English learning among the participants, but he had extensive writing experience at the research site. In addition, a year of graduate study provided him with sufficient content knowledge in his field.

Nick loved computers because “it’s my job.” Nick did not have Internet access at home. It did not seem to be a problem for his corpus use because he spent a lot of time in the engineering laboratory, where he had a computer accessible to the Internet. As in the prior study, he was positive and enthusiastic toward corpus use. He even mentioned that

⁷ As in Wol’s case, Nick took the TOEFL and GRE one year before the start of this study.
the corpus was extremely helpful for improving his performance at work. He also benefited from its portability when he later took on a summer internship with a company in California. During that time, he always used the corpus for help in writing reports, which were used for evaluating his job performance.

**Xiaodong**

The last participant, Xiaodong, 29, was a male Chinese doctoral student in molecular genetics. Xiaodong earned his bachelor’s and master’s degrees in agronomy, a field related to genetics, in China.

Xiaodong was an assiduous student who kept a busy schedule with coursework and research during his first quarter at the research site. His tight schedule made it difficult to interview him for this study, and we sometimes met only briefly before he rushed to his classes or laboratory. In fact, his writing and corpus use activities were restricted due to his lack of time.

Xiaodong had studied English “more than 10 years” from primary school to college. He learned British English because China’s English education was based on that variety. The foci of his English learning were similar to those of the others: grammar and reading, with some listening. Speaking and writing were not integrated into the school curriculum. Yet, in the first interview, Xiaodong reported that his spoken English improved very quickly after he arrived in the United States.

Xiaodong had extensive experiences in Chinese academic writing. He wrote a thesis in partial fulfillment of college graduation requirements, research papers during his
master’s degree program, and a research thesis in partial fulfillment of the requirements for his master’s degree.

With respect to L2 writing, Xiaodong’s experience was very similar to Nick’s. Xiaodong did not enjoy writing, but he liked informal writing, such as web-based chatting and e-mailing. That is, he considered these to be enjoyable communicative activities, compared to academic writing, which he found to be a strenuous activity that required much work. However, like Nick, he displayed confidence in L2 academic writing, but only when he had enough time to write. He remarked, “if I have sufficient time, I think I will be confident because I can revise.” While Nick’s competence originated from the composition classes he took in the United States, Xiaodong established his confidence by his successful L2 writing experience in his country. Notably, he published a paper in English in an international journal when he was in China. It was a remarkable achievement given that he had no publication in Chinese, and his English writing experience was limited to translation of texts and composition of short paragraphs. The paper was an offshoot of his master’s thesis, and his advisor, who received his Ph.D. degree from an American institution, helped him with revision. Another explanation for his L2 writing confidence was his high level of content knowledge, as he had pursued the same major field throughout his entire academic career.

Xiaodong liked to use computers. For him, the computer was a medium through which he could participate in many kinds of “interesting and informative” tasks, such as web-based chatting, e-mail, and Internet searches, and most of all, he found it very
convenient to use. He definitely saw the value of the corpus for his writing. He mentioned:

I understand the purpose of using it. [It’s] because we are international students. There are many steps for the word. For example, people use the word [the way] we are not familiar with. But corpus shows it to us. We will use it very conveniently. (Interview 1, 10/23/03)

Although he believed in the benefits of the corpus, he did not use it frequently because 1) he did not have time, and 2) he had no computer at home at the beginning of the study. He planned to buy one in the near future, but for the time being, he had to go to his laboratory to access the corpus.

In summary, this part of the section introduced the six focal students by delineating their salient backgrounds, past English learning, and writing experiences. It was apparent, from the beginning of the study, that the students’ L2 writing praxis was an intricate picture of their L1 academic literacy experiences, L2 learning history, disciplinary content knowledge, and educational/cultural learning contexts. It was also interesting to see that some people had a positive attitude toward corpus use while others did not at the start of the study. Why did they have different impressions and attitudes early on? What is the set of characteristics that make the use of the new method beneficial for some and questionable for others? Whether and how would their initial attitudes change over the course of time? Would those initial attitudes result in differences in their later perspectives and experiences in L2 writing? These were the guiding questions in relation to the students’ corpus use, which was the main focus of the study.
The remaining part of this chapter describes the participants’ views on language and language learning, and their writing experiences in detail. Their perspectives and experiences are reviewed collectively to synthesize common themes and issues arising from different cases, while also foregrounding each student’s salient features. Table 4.2 presents an overview of the findings, which can also be used as a guideline in reading the following discussion.
<table>
<thead>
<tr>
<th>Name (age, gender)</th>
<th>Wol (28, F)</th>
<th>Chan (29, M)</th>
<th>June (24, F)</th>
<th>Sung (32, M)</th>
<th>Nick (29, M)</th>
<th>Xiaodong (29, M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Views on language learning</td>
<td>Integrative, functional</td>
<td>Instrumental, academic</td>
<td>Integrative, functional</td>
<td>Instrumental, academic</td>
<td>Integrative, functional</td>
<td>Instrumental, academic</td>
</tr>
<tr>
<td>Views on language</td>
<td>Lexis-focused</td>
<td>Grammar-focused</td>
<td>Lexis-grammar inseparable, while Xiaodong tended toward more emphasis on grammar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence in grammar?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoy writing?</td>
<td>Yes</td>
<td>No</td>
<td>No (but enjoys informal writing)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 academic writing experience</td>
<td>Experienced</td>
<td>Very experienced</td>
<td>Semi-experienced</td>
<td>Inexperienced</td>
<td>Experienced</td>
<td></td>
</tr>
<tr>
<td>Confidence in L2 writing?</td>
<td>Yes</td>
<td>No</td>
<td>Yes (very high)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uses of L1 in L2 writing</td>
<td>Used L1 mainly for organizing</td>
<td>Thought in L1 first even for expressions</td>
<td>Rarely used L1</td>
<td>Used L1 mainly for organizing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictionary use</td>
<td>Mainly used a standard English dictionary</td>
<td>Mainly used a bilingual dictionary</td>
<td>Did not use a dictionary a lot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulties in L2 writing</td>
<td>Idiomatic expressions, language</td>
<td>Organizational and rhetorical concerns</td>
<td>No major difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 writing processes</td>
<td>Careful drafting/writing, local revision</td>
<td>Whole composition, global revision</td>
<td>Little revision due to lack of time</td>
<td>Careful drafting/writing, local revision</td>
<td>Whole composition, global revision</td>
<td>Little revision due to lack of time</td>
</tr>
<tr>
<td>Writing foci / goals of L2 writing</td>
<td>• Genuine interest in writing • Seeking for good writing beyond communication</td>
<td>• Get their ideas on paper</td>
<td>• “Good enough” writing • Communicating content</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Overview of participants’ views on language learning and writing experiences
4.2 Views on language and language learning

This section describes the students’ views on language and language learning. Given that the corpus approach is often seen as a new approach in language and language learning, the students’ initial views were examined for a comparison with their later perspectives.

Language learning is for communication

All of the participants had learned English mainly by traditional grammar-translation methods that focused on grammar and reading rather than on communication. However, they agreed that language learning is primarily for communication. They believed that learning a language is to learn how to communicate with others in either a verbal or written format. While they shared the same motive for language learning, however, their views diverged with respect to their own purpose of learning “English” (vis-à-vis a general language) and cultural learning through language acquisition.

Three participants, Chan, Sung, and Xiaodong, perceived English as a medium for communication that was necessary for achieving their academic goals (i.e., writing papers and dissertations). For them, cultural acquisition was irrelevant, and language learning was purely for academic and instrumental purposes. In contrast, the three other participants, Wol, June, and Nick, frequently expressed their genuine interest in non-academic communication and in American culture as well. They did acknowledge the importance of English in their academic studies, but they still seemed to hold the view of language as a cultural artifact. For them, language learning naturally included acquiring daily conversation skills and learning about the culture. In this sense, these three
students’ language learning can be seen as having both functional and integrative purposes. One possible explanation for this dichotomy between the former and the latter is that the students appeared to hold different views of language learning depending on their experiences and needs.

Chan and Xiaodong believed that language learning is for interacting with people who speak that language, and that it is not necessarily related to learning the culture of the speakers of the language. In addition, Chan maintained that, for him, the productive skills of speaking and writing were more important than the receptive skills of listening and reading to fully express himself in classroom discussions and academic papers. This may explain his active English learning/using style, which was seen throughout the analysis.

For me, learning English is not connected to learning the culture. Understanding culture may be important, but it’s not in my concept. For me, learning English is more like learning ways to communicate with people of the language, regardless of spoken or written form. In addition, listening and reading skills are important, but for me, speaking and writing components are much more important for now. (Interview 1, 10/28/03)

Sung generally shared the view of language learning held by Chan and Xiaodong. On the other hand, in his words, language learning is “learning grammar to transmit ideas, and also learning expressions and then ways of conveying.” Here it is interesting to note that he used the word “transmit” instead of “communicate,” which the others used. It became evident later that he was more concerned about getting his ideas across to other people than in communication. Also, it is important to note that “grammar” was
the first component of language he mentioned, which illustrates his emphasis on grammar as explained earlier in his background section.

The other participants, Wol, June and Nick, regarded the use of language in a broader context beyond academic functionality. Wol’s view was established through her experience in traveling the world, leading her to recognize the need for English as a tool for communication. Accordingly, her first year in America was devoted to improving conversation skills and learning about the new environment. Nick also devoted some of his time during his stay in America to other learning experiences than his schoolwork. Television shows were a good resource for him to learn practical English and American culture. Possibly, Wol and Nick were originally interested in the cultural aspects of the language and sought non-academic sites for learning opportunities. It was also possible, however, that they acquired their viewpoints through more exposure to communicative situations during the year preceding the study.

June had a firm conviction that language learning is integrated with culture, which is embedded in the language. In her view, language should be acquired in a natural environment, not through formal and explicit instruction, and preferably, from as early as possible in childhood.

In short, all the participants embraced the communicative nature of language, regardless of whether they recognized a connection between language and cultural acquisition. Yet, their past English learning experiences centered around grammar and reading while minimizing, if not ignoring, communicative components of the language. It came as no surprise, then, that the students expressed dissatisfaction with their learning experiences, especially those that happened before college. They provided a resounding
“no” to the question regarding whether their English education was effective in preparing them for everyday interactive communication in real-life contexts.

Sung was the lone exception. He reported that the grammar-focused instruction in secondary school prepared him with basic knowledge about the language, and that basic reading was good enough preparation for sophisticated and discourse-level reading. On the other hand, he wished he had continued to have more grammar instruction in college, so he could have performed better on the TOEFL and become a better writer in his course of the study in the United States.

In summary, the participants were first taught English by the traditional grammar-translation methods, which are structurally oriented. Yet they shared a view of learning language for communication. As ESL students, they realized the need for English as a tool for communication in their everyday life as well as in their academic studies.

**Views on language**

The students also reported their general views on language. In particular, they were asked about their perspectives on two language constituents (i.e., grammar and words) that are essential to the nature/study of language.

Sung and June represented a grammar-based view of language. Sung had a strong opinion about grammar-based instruction, as explained earlier. He had a binary concept of grammar and words; words are for representing meaning, while the purpose of grammar is to connect those words. But here, grammar, as rules, conveys ideas by putting words in the proper places. In other words, words exist discretely, and grammar
binds them together. So, for Sung, the use of prepositions (e.g., “influence” + “on”) belongs in the category of grammar knowledge, not word knowledge.

June shared Sung’s view that words and grammar are separate, and grammar plays a more important role in language construction. On the other hand, June differentiated the role of grammar in spoken and written genres. June believed that grammar is particularly important for formal writing. For her, everything in written form is grammar:

Grammar is a social agreement, and we need it for formal documents, but not for daily conversations. When we speak, we do not think of grammar. We just speak without much thinking about grammar. But when we write, we need to think about grammar because writing is something we share with others, and grammar plays the most important role in there. (Interview 2, 11/3/03)

In addition, while June claimed a clear distinction between words and grammar (“words are words, and grammar is grammar”), her concept of grammar seemed broadly applied. She stated that when more than one word is used, it is all grammar because we need to order words in the sentence. Then, given that most words do not exist by themselves, her view can be interpreted as believing that everything is grammar.

Xiaodong also placed a little more emphasis on grammar. He believed that words are given, and it is grammar that puts words in a particular order in order to make them function as language. At the same time, however, he recognized the importance of words because “without words, [there is] no grammar.” He argued that grammar is the common way people use the language. Linguists refer to this view as “descriptive grammar” as opposed to “prescriptive grammar.” He commented that “grammar is the style of what
almost other people think is right, and we have to write and speak in this way.” In this respect, he believed that grammar and words are inseparable.

Chan and Wol reflected a lexis-focused view of the language. Chan asserted that explicit grammar instruction is unnecessary beyond a certain basic level because grammar rules do not always apply in daily interactions with native speakers, and it cannot predict the way people commonly use the language. Rather, he, like Xiaodong, argued for the view of “descriptive grammar”: common use is the most representative form of grammar. Chan remarked:

I have long believed that we don’t need to learn all the grammar above a basic level. I believe grammar is represented as the form of whether people are familiar with the expressions, before they even know the grammar rules explicitly. After all, those expressions that many people use are grammatically correct, while those people do not often use are ungrammatical. Grammar can help to organize things, but it can’t explain everything. On the contrary, there are so many things that can’t be explained by grammar. So, just to follow the most common way people use the language is much more convenient. (Interview 2, 11/10/03)

In this sense, grammar and word expressions are not separable for him, while he emphasized word usage in language learning.

Like Chan, Wol stressed lexis over grammar in the nature of language and in language learning. She defined grammar as a “general, unspecified structure or rule.” For Wol, grammatical errors usually do not distort meaning, while word usage errors often result in confusion in meaning. She believed word knowledge includes the usage of so-called function words, such as prepositions. For example, that the word “influence” is followed by the preposition “on” is part of knowing the word, not a function of prepositional usage as grammatical knowledge, as some participants believed. For her,
incorrect use of prepositions (for example, “in” rather than “on” after “influence”) may sound strange, but it can still convey the purported meaning. Thus, grammar is necessary for more accuracy. She confessed that as a non-native, she could find grammatical errors rather easily, but incorrect use of words was much more difficult to notice.

Nick held the view that grammar and lexis are inseparable. It is important to reiterate that Nick had previously been exposed to the corpus for three quarters through the composition program and his own use during the summer internship. He believed that “[word] usage is the glue between vocabulary and grammar”:

At the beginning, [the grammar] were some rules, but I learn here it’s not only rules, but it’s usage. I know general rules, but now I want to learn usage than grammar. …. I think now when I start in Ohio English classes to learn about the usage, I think usage is the glue between vocabulary and grammar, but the another part I learned before is no. Vocabulary [was] vocabulary and grammar [was] grammar. (Interview 2, 11/5/03)

For Nick, the distinction between words and grammar was unnecessary because the two function as one in real language use.

To summarize, although the participants were taught by grammar-translation methods, many of them were aware of the need for an emphasis on words, or “the lexical approach.” But there is a variation in the degree to which their emphasis relates to their learning experiences and other contextual factors. On the other hand, there was no single agreement among them about the definitions and functions of grammar and words. While some dichotomized grammar and words, the others perceived them as an inseparable combination. This begs the question as to how their experience in the new corpus approach to language pedagogy would change their initial understanding of
language. I devised a follow-up question to delve into exactly that. Results will be discussed later in this dissertation.

4.3 Student approaches to L2 writing

This section describes the participants’ self-reports on their approaches to L2 writing and their perceptions of issues that inhibit/promote effective learning of L2 writing.

Importance of L2 writing: Reasons

With the exception of Nick, the other students were in their first academic term in an American university. Therefore, many of them were engaged in English academic writing for the first time. In this sense, their writing partly required “academic discourse familiarization,” as writing is an integral part of graduate-level study, with different tendencies across disciplines. Not surprisingly, then, all of the participants had one thing in common: they were interested in improving their L2 writing skills for their performance in their disciplines. They reported that skill in writing was very important for them to successfully finish coursework and write their thesis/dissertation, and finally receive their degrees. Some even ranked writing above speaking skills. For example, Sung said, “writing is of the absolute importance for getting a degree, as opposed to speaking, which is not critical for achieving my degree.”

On the other hand, the participants showed subtle differences in how they valued L2 writing. While June, Nick, and Xiaodong showed interest in developing writing aptitude, they considered it a means, not an end, for communicating their content to
others. In other words, as long as they could get their points across to others, the writing was “good-enough.” June said:

I want to be a good writer, but it doesn’t mean I am interested in writing itself. I just don’t want to spend more time in writing because of my poor skills. So, in order to focus on my [disciplinary] content, I want to be good at writing, so I don’t have to worry about writing itself and I can spend more time on the content. (Interview 1, 10/15/03)

Nick and Xiaodong expressed the same goal: they wrote only for the audience (i.e., advisors and colleagues) in their disciplines. Nick commented, “I have to write a report. If they don’t understand my writing, they don’t know what I’m doing.” Xiaodong also remarked, “you need to let other peoples know what you are doing. So writing is important. If you cannot write it out, so how do others know what you are doing?” Interestingly enough, however, both of them responded that they did not enjoy writing; they just wrote because they had to do it. Here it is worth noting that June, Nick, and Xiaodong were all in the sciences, where academic writing is considered a somewhat restricted register in terms of style and usage.

In contrast, Chan and Wol showed genuine interest in writing itself. They did not seem to be satisfied with “good-enough writing.” They wanted to not only convey meaning in their narrow academic areas, but also they wanted to be able to expound their views well in various kinds of writing contexts. Chan hoped to develop writing skills that were more expressive, creative, and intriguing, beyond basic conveyance, especially because he had changed his major focus from a scientific field to science education:

When I was majoring in chemistry in college, that kind of science field had a certain form, so it was easier to write. In contrast, now I feel I should express
myself much more clearly and persuasively than in technical writing in science. So, I feel much more burden on writing now. (Interview 1, 10/28/03)

Also, for Wol, writing was important because she could express her thoughts in a more systematic way and in more detail than in speaking. But conveying meaning was not her end goal in writing; style was something that she wanted to develop. She hoped to keep her own voice, or “writer identity,” and produce “colorful writing” as a distinct writing style, as she could do in Korean.

**Activities participants do to improve writing skills**

In order to enhance their writing skills and improve performance, the participants had preferred self-help activities. Many of them hoped to improve by “reading for writing.” As most of them were newcomers in the L2 academic setting and in a particular L2 discourse community, they believed that they needed to read more to become familiar with the target language writing conventions. Most of the students paid particular attention to good expressions while reading. Even Chan, who was described earlier as more concerned about productive skills in general and writing skill in particular, was an avid reader. But for him, reading was a way to access more resourceful and creative writing for self-study.

Taking notes and copying good sentences from reading were one of the common strategies most of the participants used. For example, Xiaodong mentioned, “when I read research papers, I do the note. There are many beautiful sentences that can express ideas very clearly. I need to write down what they have done.” It shows that the students were able to figure out, while reading, which expressions or styles were recurring and
important enough for them to follow. As the students copied good samples while reading, some (Wol and Sung) organized them in a notebook, while the others (Chan, June, and Xiaodong) did not arrange notes in a systematic way.

On the other hand, many participants (Chan, Wol, June and Xiaodong) agreed that just keeping notes was not helpful for good writing. Rather, they needed to employ the samples in their actual writing. Wol read widely from newspapers to academic texts to expose herself to various English texts and copied samples as she read. But she was unable to internalize those styles unless she applied them in her own writing. June also employed various strategies in an effort to improve writing skills: listening to news, reading essays, marking good expressions, and making notes from good sentences. But, she reiterated the need for writing practice. She mentioned that just looking at samples was not enough; she had to practice with them through her own writing.

As can be seen, the participants were active learners who were aware of their perceived weaknesses in their literacy skills and developed their own ways to compensate for their problems. In many cases, they adopted their reading selection to meet their writing needs by engaging in “reading for writing,” or broadly, trying to achieve “reading-writing connections” (Belcher & Hirvela, 2001).

**Students’ perceptions of L2 writing errors**

While the participants showed concern for precision and accuracy, they were not preoccupied with their errors in writing. As L2 learners, they expected to commit errors. In fact, most of them did not seek help to correct their writing before submitting papers to either their content courses or the writing course. The writing course required students to
submit multiple drafts, each of which the instructor reviewed, corrected and provided guiding comments on. Due to the nature of the writing course, none of the participants felt the need to polish up their papers. Neither did they get outside help when writing papers for content courses. The primary reasons were: 1) lack of time for proofreading before deadlines, 2) they were not concerned about errors as L2 learners, or 3) they were unable to locate sources for help. Xiaodong and June usually did not have time to ask for help in proofreading, as they barely met the deadlines for assignments. On the other hand, Chan and Nick tended to finish their papers before the deadlines, but they also did not try to proofread. They said their advisors often read and checked their writing for grammar as well as content. Nick said, “usually my advisor [checks] my writing. He provides sometimes grammar, but sometimes technical meaning, content, not only grammar, but content.” They believed that their teachers were good resources for proofreading because the teachers already understood they could make mistakes as non-native speakers, and they could also help with content considerations.

Wol’s most frequent errors concerned article usage. Observing her teachers’ comments, she found that even native speakers did not agree about the use of articles. So she believed that it was not a serious problem for her as a non-native speaker. At the same time, she felt that she kept making the same mistakes, such as singular/plural and countable/uncountable forms, because she did not get corrections about them:

I tend to make the same mistakes as I made before. Maybe it’s because either I didn’t make a note of the mistakes, or I was in a hurry when I turned in my paper. I think it will be very helpful if somebody can revise my writing and I keep the notes of the correction. (Interview 1, 10/14/03)
Sung was the only participant who sought outside proofreading help. In this regard, he was rather intolerant of errors compared to the others. He joined a Friday Bible study group, and one of the American group members offered to read his papers. The proofreader focused mainly on grammatical errors, such as singular-plural, prepositions, articles, and so on. This reflects not only Sung’s interest in L2 learning and writing, but also his concern about grammatical accuracy.

**Uses of L1 in L2 writing**

The participants actively drew upon their native languages in order to guide them in their L2 writing. Many of them used their L1 to plan a strategy in their L2 writing. Before they started writing, they organized their ideas, especially for complicated thoughts, in their L1. In addition, they mentally constructed phrases or sentences first in their L1s and then translated them into English. However, these were all cognitive experiences, because they did not actually write in their L1s on paper, except for unknown expressions that were first written in the L1 for easy recall and eventually replaced with approximations in English.

Chan reported that he could finish short or simple papers by thinking them out in English, while more complicated and long papers required him to think first in Korean. He said, “When I have a good flow of ideas, I just write in English. But when I have to organize my thoughts, I use Korean.” June, Sung and Xiaodong generally thought in their L1s first and started writing in English by translation. Wol thought in Korean for her L2 writing, especially for generating paper outlines, but she felt that she was gradually moving to more English. She said:
Before, I used to have an outline of the paper all in Korean. Then I translated it into English. But now I think I am in the transition stage. I mixed Korean with English. When English comes to my mind first, I just write it down in English, so I don’t forget. As a result, now my outlines look more like half in Korean and half in English. I think this is the process of change. I hope so. (Interview 1, 10/14/03)

In this regard, Nick seemed to have already made the transition, and now he felt more comfortable writing in English than in Romanian. He responded:

No, I don’t use [Romanian] now. At the beginning, yes! I made some mixture, but now I prefer to write in English, [even] take notes in English. (Interview 1, 10/24/03)

It appears that the participants first relied heavily upon their L1s for organizing ideas and outlining papers. As they became more experienced in writing and accustomed to English, they tended to think less in their L1s for their L2 writing.

**Dictionary use in writing**

Considering that the participants often translated their work from their L1s, it was not surprising that many of them used a bilingual dictionary in their L2 writing, whether it was used substantially or as a minor aid. However, they differed greatly regarding their choices of dictionaries, their satisfaction with the dictionaries, and the purpose of dictionary use.

Only Xiaodong used a paper-format dictionary, while the others used an electronic form (computer program, handheld, or Internet dictionary). Xiaodong did not often use a dictionary, but he sometimes used an English-Chinese bilingual dictionary for finding the appropriate word in a particular context. Sung also used a bilingual dictionary, which was a Korean-English handheld dictionary. He believed his dictionary
had sufficient examples for his writing. He primarily used a dictionary to find word usage.

Wol mainly used a Korean-English handheld dictionary due to its portability. But she found it insufficient in terms of content and examples. So she came to use a standard English dictionary, because it provided much better word nuances than a bilingual dictionary. Wol mainly used the dictionary to obtain the exact meaning of words, and to check nuances and accurate usage. June also preferred to use a standard English dictionary because of its numerous examples. She used various kinds of electronic dictionaries (i.e., handheld, computer program, and Internet dictionary) depending on their accessibility at the time of writing. When she encountered unknown words in a standard English dictionary, she would refer to an English-Korean dictionary to check the meaning. But she generally used a dictionary for learning the usage of words, rather than for their meanings.

Nick mainly used an online standard English dictionary, yet he sometimes used a Romanian-English dictionary when he did not understand the meaning of words. He said, “I know almost of meaning but sometimes I can't understand word very well, but very rare and not so often.” Notably, he, as a relatively long-time corpus user, emphasized the benefit of the corpus over a dictionary. According to him, “the problem is becoming more connection between phrases. Dictionary is not helpful for that. Maybe dictionary is helpful but corpus is more easy, more comfortable.”

Chan mainly used a standard English dictionary installed on his computer, while he sometimes used Internet dictionaries when available. He complained that his dictionary did not have enough examples. He used the dictionary to learn “how to
connect words, especially with prepositions, or to create phrases.” He did not use the dictionary for reading word meanings because, if he was not familiar with the words, he instead used other words whose meaning he knew.

In sum, most of the participants preferred an electronic dictionary to a paper dictionary for convenience and portability. The greatest dissatisfaction with dictionaries in general and bilingual dictionaries in particular was insufficient supply of examples in both number and coverage, while they, as advanced L2 learners, mainly used a dictionary to learn usage rather than meaning of words.

4.4 Difficulties in L2 writing

The participants reported that their main difficulties in L2 writing were, in the order of frequency of their responses: 1) insufficient command of the target language, 2) organizational and rhetorical concerns, 3) inadequate content knowledge, and 4) unfamiliar writing culture.

**Insufficient command of the target language**

Having a good command of the language was the greatest challenge that most of the participants reported facing in L2 writing. They were concerned about lexical choices, a variety of expressions, and grammar knowledge, which was seen as sentence-level language choice. They seemed to lack knowledge about exact meanings and nuances of words, usages in contexts, and their formal substitutes.

Wol had difficulty in selecting from a wide repertoire of language choices to express her ideas. She said:
I can’t fully express what I am thinking because I can’t use various vocabulary or sentence structures. I tend to use familiar expressions over and over again. So all my sentences look very similar and they are not expanding. (Interview 1, 10/14/03)

However, Wol was confident in her L2 writing, as she believed she could always generate many ideas with ease. She just wanted to have an expressive and “colorful” writing style as she did in Korean, as mentioned earlier.

Sung and Xiaodong had difficulty in choosing specific words. Sung was often unsure which word, among similar meaning-carrying words, was appropriate in a sentence. As he thought in Korean first, he struggled to express the idea in English. Xiaodong did not claim any serious difficulty in L2 writing, but reported a concern with expressing his ideas concisely and with fewer words. It appeared that he had difficulty in making appropriate academic and formal word choices in writing.

Uh, I cannot express easily. I can’t express my thought with very simple sentence. For example, sometimes Americans, they can express the meaning by using two words, but I have to use many, many words to express. I think it's the most difficult. (Interview 1, 10/23/03)

While the others were concerned about lexical choices and expressions, only June reported problems with grammar. The course teacher often corrected her article usage, so she was highly self-conscious about grammar. However, her biggest problem was organizing her thoughts and arranging them in a logical structure in her writing, which will be discussed next.
Organizational and rhetorical concerns

June expressed difficulties in overall logical organization as well as in linking paragraphs. Overall, she lacked confidence in expository writing, although she appreciated the purposes of different types of writing:

Having a logical structure is the most difficult in writing. Writing has its own style according to its type or genre, for example, simple writing, essays, or critical reviews. Simple writing is fine, but I often lack logic in the types of logical writing. I have some ideas, but in English, I can’t connect paragraphs in a way they can play their own roles in the whole organization. This is the most serious problem in my writing, although everything is problem in my writing. Well, overall, I am a poor writer. …. I don’t have confidence in writing when I have to address my opinions. I am afraid that my writing can be interpreted differently because I don’t express my ideas clearly. (Interview 1, 10/15/03)

Sung had difficulty in using discourse markers (“syntax cues” was his term), such as transitional words and phrases linking paragraphs and ideas. Wol also commented about the difficulty in linking paragraphs together. She said:

That I can’t express my ideas very well is a problem in phrase- or sentence-level expressions, but I think it can also be a problem in organization. That is, it’s because preceding ideas do not lead to the next naturally, so the paragraphs are not tied to each other. (Interview 1, 10/14/03)

Chan showed sophisticated awareness of the rhetorical features of different genres and the importance of fixed (formulaic) expressions in writing. He said the language aspect made English writing difficult, but he understood the importance of rhetorical conventions in specific genre writing. In this regard, he perceived that the nature of L2 writing is not so different from L1 writing. On the contrary, he argued that writing in Korean and English involved the same processes and difficulties. The only issue is
developing the command of language necessary to express content well. According to him:

I think organization or planning is the most difficult part in writing, either between paragraphs or within a paragraph. So when I do planning, I often use Korean to organize my thoughts. But I think writing is not different in nature whether it would be in Korean or English. I have the same difficulty in writing in Korean as in English, except for the language issue. … I think we need a lot of familiar expressions at our command to write. When writing in Korean, I have an extensive “expression bank” from which I can withdraw various expressions I need. But the bank is so limited in English. So when I try to express something in English, I always need to find how others do it. For example, when I want to report the result of research, I need to find how people express it in English. On the other hand, it seems similar even in Korean writing. When I wrote a thesis, I looked at several sample theses to see the way they wrote. Thesis writing has its own conventions and styles we need to follow. So I think the process is the same in L1 and L2 writing, except that Korean writing takes less time because I am familiar with the language. (Interview 1, 10/28/03)

These comments show that the participants’ learning of L2 writing not only involved a process of linguistic adaptation to a new environment, but was also a process of acquiring knowledge of writing conventions and rhetorics of specific genres within discourse communities.

**Inadequate content knowledge**

Overall, Nick claimed he did not have any serious difficulty in L2 writing. He was quite a confident writer. His “only” concern was “whether I know what I want to say.” In other words, as long as he had ideas about the content, he had few problems in writing it. As noted earlier, his confidence in writing stemmed from training in writing courses, which equipped him with knowledge about “structure” or “form” (i.e., sentence
construction). At this point, all he needed to be a purposeful writer was content knowledge to put in writing.

Wol reported a similar writing difficulty because she had just started her studies in a new field, so her content knowledge of her discipline was limited. June, in contrast, was pursuing the same major as in her undergraduate study. She had difficulty generating ideas and following the logic necessary to express her ideas well. However, she acknowledged that her logic problem was also related to her limited knowledge about the content. She explained that good writing requires background knowledge and logic beyond linguistic conventions:

> We need to have enough background knowledge and clear logical structure to write. For me, writing even in Korean is difficult to express my opinions. It may take me more time to write in English than in Korean, but it will take about the same amount of time in planning and organizing my thoughts in both languages. So I believe content knowledge and logical structure are the most important for good writing. … I am afraid that expert in my field would read my writing because I can be wrong and I may come to a wrong conclusion. (Interview 3, 11/19/03)

In this regard, again, the participants’ struggles can be categorized into two distinct areas: linguistic adaptation in writing and academic discourse familiarization in their disciplines.

**Unfamiliar writing culture**

Xiaodong commented about the differences between the Chinese and English writing cultures, a cause of difficulty in his writing:

> The structure is not the same when I’m writing in [Chinese] and English. The most important, greatest difference in English is you need to express your idea
and then you give the reason. In China, we usually give the reason and give the idea. I think it’s the greatest difference. (Interview 1, 10/23/03)

On the other hand, for Xiaodong, “the structure is easy to grasp because there are examples in other papers.” Because he was aware of writing differences related to culture, he did not have serious problems with the differences. He was, however, still having difficulty in acquiring detailed expressions and word usage, as noted earlier. Xiaodong’s situation recalls Nick’s comment that he could just follow structures he learned in writing courses. Therefore, he did not have problems organizing his papers.

Sung, who experienced writing in three different languages, also recognized differences in writing culture:

English seems to be the most academic language for me and difficult to write in, while Korean and Japanese are very similar. For Korean and Japanese, I only need to be careful about some different expressions, but English seems to have its own way of organization and use fixed expressions way more than the other two languages. (Interview 1, 10/21/03)

These sorts of statements demonstrate that the students, as L2 learners and writers, were acute observers of rhetorical contrasts between different cultures through their multilingual writing experiences. They applied this knowledge when approaching L2 writing tasks.

Here it became obvious that these students brought a sophisticated multilingual understanding of writing to their L2 writing experiences. They were not simply the proverbial “tabula rasa” onto which the teacher tried to implant knowledge. They consciously, or perhaps unconsciously, used other languages as frames of reference to shape their L2 writing.
In summary, the participants had various kinds of writing difficulties. It was apparent that the participants’ L2 writing involved more than a language struggle: it embraced a wide array of concerns from lexical and rhetorical language issues, to disciplinary-specific writing and content knowledge, and to writing conventions in different languages. Although the difficulties were presented in separate aspects, however, it should be noted that they often overlapped and intertwined with each other. For example, if someone has sufficient content knowledge, they are apt to be more aware of commonly used expressions and rhetoric in their disciplines. So, more often than not, difficulty in one aspect is related to one in another aspect. For analytical purposes, we must necessarily separate out these aspects. But for research to be effective, the next step would be to evaluate their interactions. This dissertation is only a first testing of these unexplored waters.

4.5 L2 writing processes

The participants were interviewed about the general processes they routinely followed in the course of L2 writing. This required them to recall the sequences and steps they followed in the actual writing of a draft that was submitted as a class assignment. Writing process research mainly employs think-aloud protocols in order to investigate students’ writing strategies and processes. In these studies, it is common to compare more proficient writers to less skilled writers, usually involving expressive writing samples. However, for this study it was considered that recall protocols were more appropriate. The participants, as advanced graduate students, engaged in disciplinary EAP (English for Academic Purposes) writing that requires extensive use of
external resources, not self-expressive writing. In this case, then, recall-protocols were more appropriate to examine participants’ writing processes.

One interesting point was that the participants compared their L1 writing to the L2 writing process in their responses. They reported similar behaviors and patterns between the two, which implied that their L1 writing or composing processes were a model for their L2 writing. Actually, this process should not be considered unusual in a cross-cultural encounter.

The salient characteristics of the participants’ writing processes are described in three stages: drafting, composing, and editing. It should be noted that the approach of the composition instructor might have affected the participants’ organizing and composing processes. The instructor’s process-oriented approach required students to write multiple drafts. This approach taught students how to construct the overall structure of the paper, so it was difficult to clearly differentiate the participants’ customary writing processes from the instructor’s approach. In fact, all of them reported that they tried to follow the structure of a paper they learned at the beginning of the course. For example, Nick reported, “I follow the instruction on schema because if I have the schema, it’s easy to fill sentences.”

On the other hand, this study attempted to examine the participants’ writing processes by comparing their first writing experiences for a paper to self-reported comments about their general writing processes at the beginning of the study.
Drafting

Most of the participants did not write an outline before they started a writing assignment. Only Wol and Sung prepared outlines for their papers. Wol wrote a detailed outline, even for Korean writing projects. As she was not familiar with English writing at the beginning, she used to write the outline in Korean. Sung also organized his thoughts in drafting his writing. Not surprisingly, then, both of them spent more time on pre-writing and planning than the others.

On the other hand, June and Xiaodong usually had a rough plan in their minds. So their planning included only an overall mental frame of writing. June said:

I have a general plan for what I am going to write in my mind, but how to start is the most difficult. I don’t know how to write the first sentence and how to make it introduce the body of content. (Interview 1, 10/15/03)

While June’s rough planning was partly related to her difficulty in generating specific ideas and organizing thoughts, for Xiaodong, time was the issue. Xiaodong devoted only a short time to pre-writing because he usually did not have enough time. According to him:

I [do not] have so many thinking before I write down. For this class, I am very busy with other things, so each time I just read the paper and have a thought in my brain and write down without thinking very carefully. (Interview 1, 10/23/03)

Chan and Nick also did not employ any extensive pre-writing techniques. Chan commonly had no written outline, unless he encountered difficulty in controlling his ideas. At that point, he tried to create one. Nick also wrote a draft only when he wanted to stimulate his idea formation:
[Drafting] depends on how well you know the problem. If I know them very well, then it is in my mind. But if I have more difficult problem[s], then I have to write something to guide, something like that. (Interview 1, 10/24/03)

In other words, only when they faced a difficult writing task or had complex ideas to organize, did they resort to writing a specific outline on paper in order to guide their writing.

**Composing**

Wol commented that she always had abundant ideas for writing assignments. Although she lacked confidence in L2 writing, she was not afraid of the composition process. She started writing after finding an organizational structure for her thoughts. Once she started, new ideas poured forth easily. She generally completed one paragraph at a time before moving on to the next paragraph, rather than moving back and forth between parts of an essay.

Sung translated his thoughts from Korean before putting them into English. He followed an outline carefully. Like Wol, he tended to complete each sentence and paragraph before moving on to the next one. Yet he revised as he wrote. He combined writing with editing, and thus, he never revised as a separate step after the first draft.

Nick also revised while he wrote. In contrast to Sung, who wrote complete paragraphs and re-organized only once, Nick wrote and revised frequently. His process was recursive in terms of writing and revising:

In my style, [I] write a sentence and each paragraph, and then secondly if I see more points, I put there. In the same manner, I do my homework. I write something like 4 pages, ok I have to write 7 pages, [so] I put more. Here I need to put a statement, and I put more data (Interview 1, 10/24/03)
I start the whole, maybe one day. [Then, if the writing] is too short, I have to put more in the background and the method. I have to make it longer, make it clearer, each part and then whole document. In the whole paper I make changes and then I see again. Oh! Maybe I have to include some more in the introduction, maybe in the conclusion, something like that. (Interview 2, 11/5/03)

Chan did not try to compose full sentences at the beginning of a writing project. He completed a rough draft of the whole paper, and then later revisited the writing, focusing on detail, resulting in a total revision. He said:

My English writing style is almost the same as my Korean writing style. I don’t make complete sentences first. I don’t care about whether it is ungrammatical or fragmented sentences first, and I just keep writing. I keep going while leaving blanks here and there. Sometimes I put only one word and move. And then I come back and make complete sentences, and after they found some shaping, I check them again as a whole (Interview 2, 11/10/03).

In the process of writing, Chan often expressed his ideas in Korean as an aid to remembering them. He believed that those translations from Korean to English resulted in awkward English at first, but he was confident he could achieve authentic English expressions eventually after repeated practice and writing.

As Xiaodong did not engage in detailed drafting, he finished the whole paper roughly all at once. After this draft, he did not make as many revisions as Nick and Chan.

Lastly, June did not complete her writing in one draft. She often made a frame or organization first for each paragraph, and then went back and “added more flesh.” June viewed writing as the process of “combining” existing phrases and text rather than creating from anew. She reported:
I copy and paste all I have, so each paragraph has several sentences in it. And then I go back and add more flesh to fill the length of requirements. …Writing is combination. It’s like combining the sentences I wrote before and the sentences from a journal article. Because I need words as a chunk in my major, if I bring those chunk words, then the sentences containing those words just come along. That’s the way I am adding a flesh to the frame. I still can’t write only with what I have in my mind. I need to remember sentence structures from reading and insert them in my writing (Interview 2, 11/5/03).

June seemed to have a hard time becoming an independent writer. She borrowed technical terms from articles and books for disciplinary content, and for general language, she used literal translations from a bilingual dictionary. She extracted phrases from the text, which she then copied in her essays. She relied heavily upon sources and was preoccupied with transmitting knowledge without reflection or interpretation.

With respect to the use of writing aids, such as a dictionary, most participants reported they made use of them when they faced a composition problem, because they could not go further without knowing the meaning of words. For example, Nick let grammar problems wait. However, he looked up the meaning of words, so he could continue on. Xiaodong was the only participant who mostly did not use a dictionary in writing. He said that he usually used alternative words instead of looking up unknown words. Yet, he added, “I think if I need to write very formal[ly], then I will use a dictionary.”

Regarding writing media, all participants except Nick wrote on the computer. They reported that a computer was more convenient and it saved time. Only Nick wrote on paper first, and then typed his hand written composition into the computer. He said, “when I write, I'm not in front of the computer, [but in a] quiet home or library, and then I type this in front of the computer. I use that to make changes.”
**Editing**

The participants reported different uses of media for different tasks at various stages. While most of the participants wrote on computer, they preferred to revise on paper. They reported that it was easier for them to revise on paper rather than on computer. For example, Nick said:

> I print it out. It’s [easier] to read from the printed out papers. It's more comfortable for me. I think and then make changes and then go to the computer, make the last modifications (Interview 2, 11/5/03).

Many of the participants engaged in revising and editing while writing, making it difficult to draw a line between where composing ended and where revision starts. They combined both into a single process, rather than into a separate “editing” stage. This illustrates the recursiveness of the writing process.

Wol and Sung did not revise their writing to a great extent, which may be related to their careful attention to each paragraph while composing a first draft. With respect to the focus of revision, Wol was mainly concerned about grammar. Wol focused on grammatical errors, such as countable/uncountable nouns and articles. She added that she rarely changed her writing extensively because she could not improve her initial expressions, and she often failed to recognize her grammatical mistakes easily. Wol said:

> The bottom line is I can’t change a lot in revision, because I put a lot of efforts when writing and that’s what I’ve finally come up with. But, after having a tutorial with the teacher, I can change with no hesitation according to his feedback. (Interview 2, 10/28/03)
In this respect, she wanted more feedback on the writing itself, including grammar, from her professors who focused primarily on the ideas and content.

Sung also said he rarely revised: “I don’t revise a lot. I don’t revise even in Korean writing. When I wrote a master’s thesis in Japanese, I also did not have revision.” This was related to the fact that he revised while writing, so he did not need any separate process in the post-writing phase. In this sense, his editing was part of his meticulous process of writing.

While Wol and Sung engaged in local editing while writing each paragraph, thus revising very little after the whole writing, Chan and Nick globally revised their assignments. They revised their writing as a whole when they finished a paper. In addition, they said they tended to revise their writing many times. Chan found that revising was self-reward after finishing a draft. In other words, it was positive reinforcement for the effort expended. Chan said:

Sometimes I revise on computer, but mostly I do on paper. It’s more like giving myself a reward for finishing a certain phase of writing. That is, I give myself a different task and mood from the one I did in front of computer. I print out the draft, walk around and look at the printouts. I move things here and there, and think about where and what to put more and write it on the paper. Then I go back to the computer and clean that up. (Interview 2, 11/10/03)

In terms of focus, Chan revised for both form and content, emphasizing content. That is, if he was satisfied with the content, then he focused on expressions and forms. If he did not like the content from the beginning, he went to work on the content.
Nick revised extensively and often. He usually finished his papers well before deadlines, so he left time to revise. This gave him time and distance for perspective. He said:

After typing to the computer, I print it out. I need to see. I need to proofread what to write more, I rewrite again, I put more extra, and I print it out, I check again, and sometimes three, four, [or] five times, I do this thing, but not in the same day, because I don't have, say, I finish my ideas in one day, and second day start to think to see where I can put this or read other papers to see, and oh! I can explain this, something like that. (Interview 2, 11/5/03)

Nick tended to focus his revisions on content. He added that his adviser often proofread his work and checked for grammar as well as content.

June and Xiaodong rarely revised globally due to lack of time. June focused on communicating content, and time permitting, she checked grammar. Her primary concern in editing was to check whether she clearly communicated her ideas on paper, and whether her writing expressed her thinking as intended. She said:

The focus of my revision is not on grammar or forms, but on content about whether this is what I want to say, or whether this is the right way to express the content. And then if time permits, I fix basic grammatical things to the extent I don’t feel ashamed of it. But I usually don’t check carefully because I have no time. (Interview 1, 10/15/03)

June’s concern about communicating content seemed to be partly related to her low confidence in writing and her worries about unskillful writing, as reported earlier.

Xiaodong did not have enough time to revise. But, if he did revise, he “concentrated most[ly] on the grammar.” He believed that revising is most critical for good writing, but his problem was lack of time. So, for him, time availability for editing
was closely related to his confidence about writing quality. He commented, “if I have a sufficient time, I think I will be confident because I can revise it.”

In summary, most of the participants’ editing was intertwined with their composition processes, regardless of whether their editing was done either locally in paragraph writing or globally after the first draft. The combined process of writing and editing may reflect their (conscious or unconscious) writing behaviors and belief that writing is a holistic process of discovery and communication of meaning. On the other hand, although the participants were at the same course level, they had different issues, problems, and approaches to writing. Some can be interpreted as characteristics of different levels of writing proficiency. We will revisit this issue later in the dissertation.

4.6 Learning styles

Another goal of this study was to identify how the participants’ initial learning styles influenced their corpus use behaviors and writing experiences. We identified learning styles through three methods: 1) self-reported comments during interviews, 2) administering a Style Analysis Survey (SAS) to each participant, and, 3) administering a Strategy Inventory for Language Learning (SILL) to each participant. Both inventories were intended to complement each other. SAS examines overall learning styles regardless of subject and language; SILL explores language learning strategies for non-native English speakers.

**Self-reported comments**

In their self-reports, many of the participants related that they tended to focus on overall understanding rather than details. Wol said she needed to understand the whole
concept first because she could not simply memorize without it. Neither did Chan try to understand every detail. Instead, he looked for overall rules while skipping minor details. Sung was also interested in getting the big picture, rather than struggling with individual details, an approach he adopted from his history studies. When faced a problem, he would skip it unless it conflicted with the overall rules. Nick reported his learning style was “to find some logic, algorithm for the problem.”

While the students were interested in finding overall rules, however, some of them had different attitudes about looking at examples. Wol needed actual examples to understand the whole concept because she could not easily understand abstractions. One of the reasons she liked her studies in education was because they often drew upon real examples to illustrate a large concept or theory. In contrast, Sung did not like a lot of examples. On the contrary, he liked concise and simple ones. As he preferred quick decisions, he said that he was personally impatient with people who hesitate.

On the other hand, June tended to focus on details as much as she valued overall rules. In terms of language learning, she was very proactive. She exposed herself as much as possible to various media, such as television news and the newspaper. She read new vocabulary and expressions out loud in order to fix them in her mind. In addition, she was a careful reader, while many of the other participants just skimmed.

The participants also related their learning styles to the types of skills required for their majors. For example, those majoring in science considered themselves to be more analytic. June viewed herself as a scientific person, lacking any artistic qualities at all. Chan was an analytic person as well. He liked interpreting quantitative data and playing with numbers. Also analytic, Nick added that he liked individual learning rather than
group learning, except when involved in a group project or if he needed to get help for a problem.

While the participants’ self reports revealed the distinct characteristics of their learning styles, the two inventories did not show much difference among them. Especially across all the sections in the SAS, they showed similar preferences or styles. As advanced graduate students, they probably shared common learning styles by necessity or from successful learning experiences that created higher education opportunities for them.

**SAS results**

Table 4.3 shows the results of the SAS. Their preferred styles are underlined.8 Most of the participants preferred visual learning styles. In other words, they learned best by seeing words and pictures in books or on handouts. Also, many of them used hands-on learning styles as well, which stressed hands-on experiences with materials. This result is consistent with Reid (1987), who said that “graduate students [in her study] indicated a significantly greater preference for visual and tactile learning than undergraduates” (p. 93). Xiaodong was the only participant whose learning style was auditory. Carriers of this style learn best from oral explanations, class discussions, lectures, or audiotapes. Reid also found that Korean students were the most visual among learners from various language backgrounds, and the Chinese, in contrast, were auditory, which was shown in the participants’ learning preferences in this study.

---

8 Students are asked to respond to each statement of five sections by using a scale of 0 to 3 (0 = never, 1 = sometimes, 2 = very often, and 3 = always). Then the scores are tallied. The largest score is considered their style preferences. If there are scores within 2 points apart, both are considered their styles.
All participants tended to be extroverted, as they liked to interact and learn with others. At the same time, Chan, Sung, and Nick also showed an introverted tendency, because they sometimes preferred individual learning. Regarding the handling of possibilities, most of the participants employed intuitive styles. According to SAS instructions, “intuitive” learning styles are “future-oriented, able to seek out the major principles of the topic, like to speculate about possibilities, enjoy abstract thinking, and avoid step-by-step instruction,” while “concrete-sequential” styles are “present-oriented and prefer one-step-at-a-time activities.” Only Nick used both intuitive and concrete-sequential learning styles, while the others primarily adopted intuitive styles.
Most of the participants approached tasks in a closure-oriented way. They preferred structured directions and planned their assignments ahead, rather than enjoying “discovery learning without concern for deadlines or rules” (SAS instructions, 1989). Wol used an “open” learning style. Regarding dealing with ideas and tasks, most of them tended to be global in orientation. That is, they sought to understand general principles rather than focusing on details. Only June reported using an analytic learning style in addition to the global style.

**SILL results**

Table 4.4 shows the SILL results. The participants’ commonly used strategies (that is, those with an average over 3.5) are underlined in the table.9

<table>
<thead>
<tr>
<th>Part</th>
<th>Strategies</th>
<th>Wol</th>
<th>Chan</th>
<th>June</th>
<th>Sung</th>
<th>Nick</th>
<th>Xiaodong</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Remembering more effectively</td>
<td>2.00</td>
<td>2.67</td>
<td>2.89</td>
<td>2.78</td>
<td>3.11</td>
<td>3.11</td>
</tr>
<tr>
<td>B</td>
<td>Using all your mental processes</td>
<td>3.79</td>
<td>2.93</td>
<td>3.43</td>
<td>3.07</td>
<td>3.43</td>
<td>3.50</td>
</tr>
<tr>
<td>C</td>
<td>Compensating for missing knowledge</td>
<td>3.50</td>
<td>3.50</td>
<td>3.67</td>
<td>2.83</td>
<td>2.67</td>
<td>4.00</td>
</tr>
<tr>
<td>D</td>
<td>Organizing and evaluating your learning</td>
<td>2.89</td>
<td>2.78</td>
<td>4.00</td>
<td>3.89</td>
<td>4.00</td>
<td>3.44</td>
</tr>
<tr>
<td>E</td>
<td>Managing your emotions</td>
<td>2.33</td>
<td>2.17</td>
<td>3.17</td>
<td>3.00</td>
<td>1.67</td>
<td>2.67</td>
</tr>
<tr>
<td>F</td>
<td>Learning with others</td>
<td>2.33</td>
<td>3.33</td>
<td>4.33</td>
<td>3.67</td>
<td>3.67</td>
<td>3.17</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>2.92</td>
<td>2.88</td>
<td>3.54</td>
<td>3.20</td>
<td>3.20</td>
<td>3.34</td>
</tr>
</tbody>
</table>

Table 4.4: Results, Strategy Inventory for Language Learning (SILL)

The table shows that June, Sung, and Nick used the same major strategies. They consciously monitored and organized their language learning while interacting with

9 The average scores show the degree of using a corresponding strategy. That is, an average of 4.5 to 5.0 = always or almost always used, 3.5 to 4.4 = usually used, 2.5 to 3.4 = sometimes used, 1.5 to 2.4 = generally not used, and 1.0 to 1.4 = never or almost never used.
others. On the other hand, the other three, Wol, Chan, and Xiaodong, generally used similar major strategies. They all tried to compensate for gaps in knowledge by applying their background knowledge or by using alternative measures. They also expended substantial mental energy into pursuing more effective language learning (e.g., trying not to translate word-for-word and finding patterns in English).

To summarize, the self-reports and inventories’ results manifested no recurrent distinct characteristics when combined. However, they did reveal distinct features of some of the participants’ learning styles. These are global vs. analytic-oriented preferences, and compensation for missing knowledge and learning with others. Also revealed were their personalities, as described in the previous sections. These revelations begged interesting follow-up questions, which were pursued in the later interviews. Would those differences in preferences result in any differences in participants’ corpus use behaviors and beliefs? Would those differences be related to their writing experiences and language learning in general? These questions will be explored in depth later in this dissertation.

4.7 Summary

This chapter introduced the six case study participants by focusing on their salient personal background characteristics, and major language learning and writing experiences. Also described were their initial views on language learning, approaches to L2 writing, difficulties in L2 writing, L2 writing processes, and learning styles. Throughout the section, it was apparent that the participants did not bring blank minds to the formal classroom learning process. Rather, these advanced L2 learners brought
unique personal histories, rich with knowledge and experiences, to the new academic environment, and they integrated them into their new learning. This is important because they were all at the same course level, but each of them had their own issues and difficulties in L2 writing that were shaped from their past learning and writing experiences. As active learners, they developed individual ways to compensate for gaps in knowledge. Most importantly, L2 writing was more than just an academic exercise in L2 learning for its own sake. It was a pathway to their own disciplinary success. Engaging in academic writing was also a critical way of acquiring disciplinary conventions, academic literacy, and entrée into a discourse community. Writing is one of the most important scholarly tasks in graduate studies. These data show they were engaging both in linguistic adaptation and in academic discourse acculturation. This can be an extremely challenging task, especially given that most of them had minimal experience in L2 writing before coming to America (e.g., only TWE writing and writing by translation).

An interesting finding is that the participants could be paired in three groups: Wol & Chan, June & Sung, and Nick & Xiaodong. The pairs showed striking similarities in their perceptions and experiences in L2 learning and writing (see Table 4.2 for a quick illustration).

The first pair, Wol and Chan, held a lexis-focused view of language. They generally enjoyed writing and were confident in L2 writing. Neither were they afraid of L2 writing, and they believed any writing was manageable. Most of all, they had a sincere, genuine interest in L2 writing. Thus, their goal in writing was not just to communicate ideas, but to develop persuasive and powerful texts. At the same time, they
wanted to develop their own writing styles. Wol even hoped to keep her unique L1 voice in L2 writing.

The second pair, June and Sung, emphasized the grammatical aspects of the language while they lacked confidence in grammar. In general, they disliked writing. Both of them were inexperienced in L1 academic writing. For them, L2 writing was a burden that required them to work hard to get across their points. Similarly, their primary concern in writing was to get their ideas across on paper, with an emphasis on linguistic features.

The last pair, Nick and Xiodong, held a balanced view between language in lexis and grammar. They believed that the two components were inseparable, and the dual concept approach should be realized in language learning. Generally, they did not like writing. They wrote only because they had to do it. However, both of them enjoyed informal writing for communicating with others. Likewise, their primary focus in writing was to communicate content to readers. They viewed the purpose of writing as making themselves understood by others. In another sense, they did not care so much about text as long as their writing was “good enough.” In addition, they showed a great level of confidence in L2 writing.

These pairings can help readers view the participants across the cases with more ease while identifying each of them. This initial pairing will be tested throughout the main analysis of the study. In addition, interesting questions regarding this pairing arise as to whether the pairs would show similarities in their corpus use behaviors and later perceptions of language and writing. Based on the questions and information provided in this section, the next section will describe the use of corpus technology in L2 writing.
CHAPTER 5

STUDENT USE OF CORPUS TECHNOLOGY IN L2 WRITING

This chapter is primarily devoted to presenting the findings from all the data sources used in the study. The next chapter, Chapter 6, will be devoted to interpreting what those findings mean as related to the study’s research questions.

The findings are presented in three sections in this chapter. Section 5.1 reports the participants’ corpus use patterns and their perspectives on its use in L2 writing. These qualitative findings not only present the unique features of the focal students’ corpus use experiences and their perceptions of its advantages and disadvantages, but they also provide in-depth analysis of its pedagogical values in L2 writing instruction. Section 5.2 presents any changes that were observed in the participants’ writing experiences and language learning perspectives after their involvement in corpus use. In particular, it describes how corpus use affects their understanding of the target language, their approaches to L2 writing, and their writing processes. The last section, 5.3, presents the role of collocation in L2 learning and writing as suggested by the participants.
5.1 Students’ corpus use in L2 writing

This section reports the findings of the six case study students regarding their corpus use experiences and evaluations of its value in L2 writing. Various kinds of qualitative data sources were employed: interviews, think-aloud protocols, student corpus search logs, weekly search assignments e-mail to the teacher, and written reflections on corpus use. The participants also participated in a class survey, and the results of the survey were incorporated into subsequent interviews. In this section I first present the participants’ corpus use behaviors in detail and then move on to their perceptions of its benefits and problems. Unfortunately, the online corpus service malfunctioned after the first quarter of the study. The technological breakdown made the interviews in the second quarter address hypothetical questions rather than real-world experiences. Questions were restructured to ask students how they would have used the corpus if it were available.

5.1.1 Corpus use patterns and benefits

Classroom contexts and activities

In order to provide a context for the study, this part of the section starts with descriptions of the classroom environment and activities related to corpus use. The students involved in this project were enrolled in an advanced level course in the ESL academic writing program. Out of a total of 14 students, nine were male and five were female. Their ages ranged from 23 to 35. In terms of nationality, Koreans dominated; there were 10 Koreans, three Chinese, and one Romanian. They were all from different
programs of study, with ten in science-related fields, two in the arts, one in education, and the other one in humanities. Nine of the students were working on doctoral degrees, while the rest were in their masters’ programs. Four students had already been in America for one year, while the remainder were new arrivals.

The class met twice per week for two and a half hours each session. Throughout the course, the students were required to perform three major writing tasks: an annotated bibliography, a critical review, and a final long paper chosen by the student among several options. The latter could be a literature review, critical essay, research proposal or report. The students’ own interests and needs in their studies determined the topic or content of the writing. In this sense, they were engaged in disciplinary EAP writing.

The first day of class the instructor introduced students to the course, including information about the syllabus, books, and course assignments. Also, an in-class writing task was conducted for diagnostic purposes. Those who did extremely well on the test were exempted from the course. On the second day, the course instructor introduced the Collins COBUILD corpus to the students. With a handout, the teacher demonstrated how to use the corpus online sampler and interpret the search results. To provide a brief introduction, the corpus was shown having two types of search output on separate pages: “concordance” and “collocate” output. The concordance function allows the student to look at an unknown word, “target word,” located in the middle of each sentence, with a number of sentences presented. The program then shows the textual context on the left and right of the word being investigated. For example, in Figure 5.1, the target word is “usage,” and each sentence displays the word in the middle with its surrounding text, showing the occurrences of the word in context.
scheme. Based on your past electricity usage, we estimate that the total of these FIVE news and current affairs and some obs. This usage has been reflected in other for shots are used they must be in common usage so that they are easily understood e.g gv Exact prices are available on request. Usage charges for outgoing service are available A Survey: on technology and information usage and needs in the voluntary sector. by building vocabulary, learning idiomatic usage and reducing your accent, you will also prepared to accept a need for restraint on car usage. Attitude surveys are fine. We know the development of contact lenses and their usage by the public. [p] [p] [h] Mechanisms amount irrespective of its size. If your water usage is very small it may be worthwhile

Figure 5.1: Example of concordance output for the word “usage”

On the other hand, the collocate output is a list of words commonly used in conjunction with the target word. This list is organized according to significance of occurrence. Upon a search of a word, the corpus presents 40 randomly chosen concordances and the most commonly used 100 collocates (see Figure 5.2 for an example). While concordances help to show how those words are actually used in sentences, collocates provide scientific data about how often the words co-occur within the corpus.
<table>
<thead>
<tr>
<th>Collocate</th>
<th>Corpus Freq</th>
<th>Joint Freq</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>in</td>
<td>765730</td>
<td>59</td>
<td>3.387999</td>
</tr>
<tr>
<td>of</td>
<td>1100578</td>
<td>77</td>
<td>3.373624</td>
</tr>
<tr>
<td>word</td>
<td>7972</td>
<td>10</td>
<td>3.053712</td>
</tr>
<tr>
<td>notes</td>
<td>2197</td>
<td>8</td>
<td>2.794976</td>
</tr>
<tr>
<td>common</td>
<td>5901</td>
<td>8</td>
<td>2.738579</td>
</tr>
<tr>
<td>this</td>
<td>190140</td>
<td>20</td>
<td>2.641152</td>
</tr>
<tr>
<td>English</td>
<td>8537</td>
<td>7</td>
<td>2.506794</td>
</tr>
<tr>
<td>water</td>
<td>13608</td>
<td>7</td>
<td>2.424253</td>
</tr>
<tr>
<td>modern</td>
<td>5310</td>
<td>6</td>
<td>2.356133</td>
</tr>
<tr>
<td>into</td>
<td>65158</td>
<td>10</td>
<td>2.274930</td>
</tr>
<tr>
<td>energy</td>
<td>4287</td>
<td>5</td>
<td>2.153503</td>
</tr>
<tr>
<td>your</td>
<td>91112</td>
<td>11</td>
<td>2.133569</td>
</tr>
<tr>
<td>such</td>
<td>31861</td>
<td>7</td>
<td>2.127147</td>
</tr>
<tr>
<td>reviewed</td>
<td>377</td>
<td>4</td>
<td>1.991882</td>
</tr>
<tr>
<td>note</td>
<td>3495</td>
<td>4</td>
<td>1.924744</td>
</tr>
<tr>
<td>charges</td>
<td>3582</td>
<td>4</td>
<td>1.922870</td>
</tr>
<tr>
<td>continued</td>
<td>3806</td>
<td>4</td>
<td>1.918047</td>
</tr>
<tr>
<td>average</td>
<td>4375</td>
<td>4</td>
<td>1.905795</td>
</tr>
<tr>
<td>with</td>
<td>290716</td>
<td>21</td>
<td>1.850546</td>
</tr>
<tr>
<td>term</td>
<td>7193</td>
<td>4</td>
<td>1.845116</td>
</tr>
<tr>
<td>low</td>
<td>7563</td>
<td>4</td>
<td>1.837149</td>
</tr>
<tr>
<td>information</td>
<td>11751</td>
<td>4</td>
<td>1.746971</td>
</tr>
</tbody>
</table>

Figure 5.2: Example of collocate output for the word “usage”

After an introduction to corpus technology, the teacher taught the class how to use concordance and collocate data to construct “prototype strings” of each word. Prototype strings are the most commonly used forms of word usage. They are different from traditionally used scientific-looking structures that often use metalanguage (e.g., “difficulty” + “generative” form) and are usually focused on the structure or form of
words. In contrast, the prototypes use actual words used to construct the representative usage patterns. After demonstrating a prototype construction on the chalkboard, the teacher asked the students to work in groups to produce prototype strings of some words. The students reported their findings to the class, while the teacher responded with comments. Figure 5.3 is an example of constructed prototypes extracted from class activities. It was constructed by using the most common collocates of “knowledge” while referring to its concordances. In fact, the teacher chose the word “knowledge” after he found in students’ writing the phrase “learn knowledge,” which is grammatical but doesn’t sound right. The constructed prototype shown in the box below shows the appropriate verbs in the first column and common modifiers in the second column. Even advanced L2 learners can find modifiers to be difficult choices because they express how to quantify or qualify abstractions.

<table>
<thead>
<tr>
<th>&lt;KNOWLEDGE&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>denied</td>
</tr>
<tr>
<td>acquire</td>
</tr>
<tr>
<td>experience</td>
</tr>
<tr>
<td>gain</td>
</tr>
<tr>
<td>advanced</td>
</tr>
<tr>
<td>detailed</td>
</tr>
<tr>
<td>specialized</td>
</tr>
<tr>
<td>working</td>
</tr>
<tr>
<td>accumulated</td>
</tr>
<tr>
<td>knowledge of language</td>
</tr>
<tr>
<td>repeatedly about</td>
</tr>
</tbody>
</table>

Figure 5.3: Example of student prototype construction

The purpose of the prototype construction activity was 1) to make students aware of the most common usage patterns, and 2) to ease the writing process with a convenient on-hand listing. The activity was based on the teacher’s belief that teaching structure or grammar directly does not help students much in their actual writing. He was convinced...
that they need to work with real words, and “usage may be more important than grammar for advanced learners.” He believed readers as well as writers could process information more quickly through this method.

The teacher wanted the students to integrate corpus use into their writing for everyday purposes. He encouraged them to identify their own writing problems so they could use the corpus to solve the problems on their own. In this sense, the corpus was used as both an instructional tool and also as a dynamic resource for independent writing. The teacher required students to search the corpus regarding their own writing problems and to e-mail the search results weekly to him. Then he combined those results on handouts regularly provided to the class so they could benefit from each other’s corpus searches. In addition, he usually began class sessions by commenting on writing errors that he found in students’ drafts. He encouraged them to research the problems through the corpus. He also wrote feedback on their papers, directing them to search out solutions rather than correcting errors immediately. By so doing, he expected that by the end of the course, the class would generate a useful lexicon that stemmed from their own errors.

At the end of the course, a survey was administered to all students in the class in order to examine the class’ overall use of corpus technology in L2 writing instruction. While this dissertation does not discuss the results of the class survey in detail, we briefly review the main points. Overall, the class students were positive about using the corpus in L2 writing. They perceived corpus use to be helpful for improving L2 writing in general and for acquiring common usage patterns in particular, which confirmed the purported usefulness of corpus integration into the writing course. Corpus use not only
supplied them with the concrete linguistic input, but it also elevated their confidence about L2 writing. They did not report any major problems in using the corpus, except that it was time-consuming. All in all, they were fairly positive about the advantages of the corpus as a useful resource for L2 writing. Summary tables of the results are provided in Appendix I.

**Corpus use: Frequency and purpose**

It may be helpful to recall the six case study participants’ initial responses to corpus use and their attitudes toward technology from the background portraits (for a quick reminder, see Table 4.1: Overview of case study participants). Four of the participants, Wol, Chan, Nick, and Xiaodong, were positive about corpus use. In particular, Nick was very positive because he believed he benefited from it after using it for one year. In contrast, June and Sung were negative toward corpus use. June was skeptical because of her difficulty in using the technology. On the other hand, Sung preferred using a traditional reference – dictionary – over the corpus. Internet accessibility was a limiting factor because only Wol and June had Internet access at home. June, however, usually did not use the home Internet service because she did not want to work at home. The Internet was unavailable at home for the other four participants. But, for Nick and Xiaodong, it was not a problem because they spent most of their time in their laboratories, where they had access to computers.

The participants’ corpus search logs, e-mail assignments, and interviews revealed their frequency of corpus use, the purpose of specific corpus searches, the frequency of searches for items (or linguistic features), and their behaviors while searching. They kept
a log of their searches, which they brought to the following interview. In the log, they recorded search items, the purpose of the search, their level of satisfaction with the search results, and the amount of time spent for the searches. The log form is attached in Appendix D.

Table 5.1 shows the participants’ comments about their purposes for using the corpus. It reflects their perceptions of the advantages of corpus use because it expresses the reasons they chose to use the corpus.

<table>
<thead>
<tr>
<th>Name</th>
<th>The purpose and benefit of corpus use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wol</td>
<td>1) preposition, 2) correct usage, 3) phrasal verb</td>
</tr>
<tr>
<td>Chan</td>
<td>1) delicate differences or nuance, 2) preposition, 3) modifiers later</td>
</tr>
<tr>
<td>June</td>
<td>1) preposition, 2) verb + noun connection, 3) nuance</td>
</tr>
<tr>
<td>Sung</td>
<td>1) preposition, 2) verb + noun connection, 3) word use context</td>
</tr>
<tr>
<td>Nick</td>
<td>1) preposition</td>
</tr>
<tr>
<td>Xiaodong</td>
<td>1) normal usage</td>
</tr>
</tbody>
</table>

Table 5.1: Participants’ purpose of corpus use

The table shows that the majority used the corpus to clarify appropriate preposition usage in sentence construction. Sung commented, “I use the corpus when I have a problem in prepositional phrases.” This problem clearly reflects on their native languages. Korean, Romanian, and Chinese\(^{10}\) do not have prepositions. Hence, they searched for help in the corpus. Nick said:

---

\(^{10}\) Dr. Jack Rouzer, a committee member specializing in Chinese Linguistics, pointed out that “it’s true that Chinese does not have prepositions, but it does have words to specify spatial relations which, as notions, are often directly translatable into English.” He went on to point out the complexity of the influence of Chinese L1 pattern existence or non-existence on L2 learning. According to him, “the fact that Chinese represents spatial relations with postpositions (in combination with one preposition) does not
I [would] say the corpus is for preposition because it’s my problem. My language [Romanian] is almost like French. 80, 90% is French. So pretty much like Latin. It’s a Latin language. Latin language doesn’t have preposition. So, preposition is my big problem. (Interview 4, 10/24/03)

Many of the participants also used the corpus to check the correct usage of words, that is, how the words are used in sentences, which words are often used together, e.g., verb and noun correlations. The participants added that the corpus data showed nuances, delicate differences of word meanings, and contexts of uses. June said, “the corpus is the most helpful for learning nuances of words that carry similar meanings, but are used only in specific contexts with only certain words.”

In terms of word usage, the participants praised the corpus mainly for finding words that are perceived to play an important role in sentence construction, such as verbs and prepositions. Chan, in sharp contrast to the others, showed interest in the usage of modifiers. Like the others, he reported at the earlier stage of the study that the corpus was beneficial for learning prepositional phrases and verb usage. As the study progressed, however, his interest changed to the usage of adjectives and adverbs. According to him:

Recently, what I look up the most in the corpus is actually adjectives and adverbs, that is, which words are the most appropriate to modify nouns or verbs. When I want to modify a “barrier,” is it “severe,” “heavy,” “great,” or “major”? I was not so much concerned about this aspect before. I just used any words. In fact, I thought that verbs, nouns and prepositions were important skeletons of sentences, but modifiers were trivial and they did not make a big difference. But I came to realize that they also play an important role in the sentences. Those modifiers explain the difficulty that Chinese learners of English may have with learning English preposition usage. On the contrary, given that analogous spatial relations are represented lexically in both languages – albeit with postpositions in Chinese – could facilitate Chinese people’s learning of English” (e-mail correspondence, Feb. 3, 2005).
contribute to making much more fluent and eloquent sentences. So, now I am concerned about them a lot. I think this is the biggest change in my writing. (Interview 2, 11/10/03)

Chan added that dictionaries cover prepositional usage very well, but are limited in the number of examples about modifiers. His gradual change of interest in linguistic features may have been due to his increased amount of writing experience and linguistic knowledge. While focusing more on substantial linguistic elements of sentence construction at the beginning, his enhanced writing skills and knowledge drew his attention to the less crucial features that are “seasoning” to the language.

In addition to the participants’ general comments, Table 5.2 shows the participants’ main corpus search items taken from search logs and e-mail assignments. The table shows the amount of corpus use varied among the participants. On the whole, Wol, Chan, and Sung used the corpus much more frequently than June, Nick, and Xiaodong, which was also confirmed through interviews. The former three used it for other courses and their own needs in L2 writing, as well as for the writing class, while the latter three reported that they only used the corpus for class writing assignments. One conspicuous difference between the two groups, which might explain this, is the field of study. The latter three were science majors, where they may have been less need for writing in their courses. On the other hand, the former three were non-science majors (i.e., education, history, and science education), all likely requiring more writing, and of a wider varieties of styles.
<table>
<thead>
<tr>
<th>Purpose</th>
<th>Chan</th>
<th>Wol</th>
<th>Sung</th>
<th>June&lt;sup&gt;11&lt;/sup&gt;</th>
<th>Nick</th>
<th>Xiaodong</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item</td>
<td>Remark</td>
<td>Helpful</td>
<td>Time&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Item</td>
<td>Remark</td>
</tr>
<tr>
<td>Proper perception</td>
<td></td>
<td></td>
<td>yes</td>
<td>5'</td>
<td>accompany</td>
<td>yes</td>
</tr>
<tr>
<td>be enrolled</td>
<td></td>
<td></td>
<td>yes</td>
<td>5'</td>
<td>occupy</td>
<td>yes</td>
</tr>
<tr>
<td>enthusiastic</td>
<td></td>
<td></td>
<td>yes</td>
<td>2'</td>
<td>collaborate</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>detail</td>
<td>yes</td>
<td>30''</td>
<td>offer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>adopt</td>
<td>yes</td>
<td>30''</td>
<td>role</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>adapt</td>
<td>yes</td>
<td>30''</td>
<td>dimension</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>apply</td>
<td>yes</td>
<td>30''</td>
<td>education</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>identity</td>
<td>yes</td>
<td>30''</td>
<td>level</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>impact</td>
<td>yes</td>
<td>30''</td>
<td>insight</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>example</td>
<td>yes</td>
<td>30''</td>
<td>identifying</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>enroll</td>
<td>yes</td>
<td>30''</td>
<td>evaluation</td>
</tr>
</tbody>
</table>

Table 5.2: Participants’ corpus search items

<sup>11</sup> June, Nick, and Xiaodong did not indicate the amount of time and the level of satisfaction with the results of corpus search in the logs.

<sup>12</sup> This column shows additional purposes or prior guesses of the search given by the participant.

<sup>13</sup> For the amount of time, ' indicates minute, and " indicates second.
Table 5.2 continued

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Chan</th>
<th>Wol</th>
<th>Sung</th>
<th>June</th>
<th>Nick</th>
<th>Xiaodong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usage/contexts of uses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regarding</td>
<td>to?</td>
<td>yes</td>
<td>8'</td>
<td>yes</td>
<td>30''</td>
<td>regard</td>
</tr>
<tr>
<td>have difficulty</td>
<td>yes</td>
<td>15'</td>
<td></td>
<td></td>
<td></td>
<td>pen(verb)</td>
</tr>
<tr>
<td>understanding (noun?/verb?)</td>
<td>yes(++)</td>
<td>30'</td>
<td></td>
<td></td>
<td></td>
<td>embed</td>
</tr>
<tr>
<td>addressing (noun?/verb?)</td>
<td>yes(++)</td>
<td>10'</td>
<td></td>
<td></td>
<td></td>
<td>evaluate</td>
</tr>
<tr>
<td>both in?/in both?</td>
<td>yes</td>
<td>7'</td>
<td></td>
<td></td>
<td></td>
<td>seem</td>
</tr>
<tr>
<td>on the contrary</td>
<td>no</td>
<td>15'</td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>to the contrary</td>
<td>no</td>
<td>30'</td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>in contrast by contrast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
</tr>
</tbody>
</table>

149 The participant marked level of satisfaction
15 Chan searched the same item three times on different days.
Table 5.2 continued

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Chan</th>
<th>Wol</th>
<th>Sung</th>
<th>June</th>
<th>Nick</th>
<th>Xiaodong</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item</td>
<td>Remark</td>
<td>Helpful</td>
<td>Time</td>
<td>Item</td>
<td>Remark</td>
</tr>
<tr>
<td>V+N¹⁶</td>
<td>review(noun)</td>
<td>summary</td>
<td>verb?</td>
<td>little</td>
<td>yes</td>
<td>10'</td>
</tr>
<tr>
<td></td>
<td>war</td>
<td>problem/issue</td>
<td>discussion</td>
<td>verb?</td>
<td>little</td>
<td>yes</td>
</tr>
<tr>
<td>V+N</td>
<td></td>
<td></td>
<td>depict</td>
<td>enunciate</td>
<td>what kind</td>
<td>object?</td>
</tr>
<tr>
<td>adj+N</td>
<td>frequency</td>
<td>high?</td>
<td>yes</td>
<td>3'</td>
<td>bias</td>
<td>mistake</td>
</tr>
<tr>
<td>adv+adj</td>
<td>lower</td>
<td>quite?</td>
<td>yes</td>
<td>2'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adv+V</td>
<td>affect</td>
<td>yes</td>
<td>20'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V+adj</td>
<td>feel</td>
<td>difficult?</td>
<td>yes</td>
<td>8'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meaning</td>
<td></td>
<td></td>
<td>suspect</td>
<td>status</td>
<td>turmoil</td>
<td>yes</td>
</tr>
<tr>
<td>Form (sg/pl)</td>
<td>process</td>
<td>countable?</td>
<td>yes</td>
<td>5'</td>
<td>threat</td>
<td>V form?</td>
</tr>
</tbody>
</table>

¹⁶ For simple representation, linguistic class terms for the search are presented in acronyms. That is, “V” stands for verb, “N” for noun, “prep” for preposition, “adj” for adjective, and “adv” for adverb. Also, target searches are in bold. For example, “V+N” indicates that the participant searched the noun in order to find appropriate verbs that are commonly used with the noun.
Regarding the frequency of corpus use, quite surprising are Sung and Nick’s reports that seemingly counter their initial attitudes to corpus. That is, Sung showed a clearly negative attitude to corpus at the beginning, but later he was one of the participants who used the corpus to a great extent. In contrast, Nick showed the most positive attitude initially due to his previous one-year exposure to corpus, but he used it minimally, mostly for the writing class assignments. It was an unexpected result because Nick was the only participant who reported that he habitually used the corpus in his writing.

With respect to search items, many of the participants most frequently searched for appropriate preposition usage, as indicated by their self-reports. The next most common use of the corpus was for correct usages and contexts of words. Such usages include questions about what kind of complements verbs take (e.g., clause or noun form), whether verbs are commonly used as active or passive forms (e.g., “consist of” vs. “is consisted of”), and which collocates the verbs take. The most frequently searched word class was verb form, which shows that the participants had the most interest or difficulty in verb usage.

A few participants used the corpus to learn about appropriate noun and verb correlations. For example, Sung searched the verb “depict” to find the kinds of nouns normally used with the verb. On the other hand, Chan was the only participant who searched for adjective and adverb usage, as mentioned earlier. For instance, he searched the word “frequency” to find which adjectives can be used to quantify it. The fact that he searched for modifiers indicates his sophistication as a corpus user.
Two participants, Sung and June, also used the corpus for checking word meanings. Sung examined numerous examples of the concordance output to clarify the most probable meanings of words in various textual contexts. June, on the other hand, used the corpus like a dictionary; she searched for word definitions in the corpus and guessed at word meanings rather than looking up examples to clarify the meanings. For example, she looked up “a priori” in the corpus, but she failed to find the meaning from the corpus search.

This study also investigated the amount of time that participants spent on searches. Participants were asked to keep a log of their searches, which recorded the time consumed as well as items searched, the purpose of the search, and the level of satisfaction with the results. Unfortunately, only three participants, Chan, Wol, and Sung, kept detailed logs and tracked the time spent on searches. The other three said they did not log their entries, thus limiting the analysis. On average, the amounts spent on searches were: Chan, 12 minutes, Wol, 43 seconds, and Sung, 3 minutes. Wol, whose searches took the least time, looked mainly for prepositions and phrasal verbs. Sung’s searches generally took more time than Wol’s. Interestingly, Sung’s search logs showed that his later searches took an average of 4 minutes, more than four times as much as his earlier ones, which often took less than one minute. A closer look at the logs revealed that his earlier searches focused mainly on prepositions, while later searches expanded into a wider repertoire of linguistic features, requiring more time than prepositional usages. This included word usage and collocation.

Chan’s corpus search time and behaviors, as shown in Table 5.2, deserve special attention. He usually spent a great deal of time on his searches and provided the most
detail in his logs. He expended a great deal of effort in analyzing corpus data; his searches and analyses were very sophisticated, persistent, and purposeful. This does not mean the other participants’ searches lacked similar intensity. They provided less detail, or else failed to keep a log, so we cannot draw conclusions. Still, Chan’s log is very useful in investigating search processes. In one search, Chan looked for the right sequence of “both in” or “in both” on three different days, spending a total of 52 minutes. His sophisticated and enthusiastic corpus analysis is well represented in another example. Figure 5.4 is extracted from one of his e-mail assignments. He searched for “problem/issue” in order to discern nuances and context of uses after the teacher pointed out his interchangeable uses of the two. His purposes in the corpus searches were two-fold: 1) to differentiate the nuances between the two words, and 2) to identify common verbs to collocate with them. He constructed prototypes for each word from the concordance and collocate output and made a comparison between the two. As seen in his conclusion, he finally identified commonly used adjectives (italics in original) as well as the verbs (bold in original) for each word. This illustrates how he solved his writing problem and, furthermore, how he expanded his linguistic knowledge by using the corpus.
<VERB+problem@/NOUN>

# can cause related problem(s)
may solve underlying (No.4 / Significance 7.97)
could create pressing (No.18 / Significance 3.84)
overcome increasing (No.19 / Significance 3.79)
tackle continuing (No.21 / Significance 3.67)
resolve

discuss pose prevent created face avoid

<VERB+issue@/NOUN>

# discuss related issue(s)
address burning (No.5 / Significance 4.63)
raised economic (No.11 / Significance 4.09)
share complicated (No.24 / Significance 2.66)
resolve underlying (No.25 / Significance 2.61)
identified raises pressing
resolved selected
identify clarify dealt with

# refund the unmailed issues

Conclusion

1. Because an ISSUE is more burning (hot or controversial) and complicated than a PROBLEM, the issue needs to be addressed, shared and discussed before being resolved. However, a PROBLEM is pressing, increasing and continuing one that has an apparently negative effect it should be solved, overcome, tackled and resolved.

2. Because an ISSUE is more complicated one, it needs to be identified and clarified; however a PROBLEM is apparently negative one, it does NOT need to be identified and clarified. Surely, a PROBLEM should NOT be shared.

3. “solve the problem” and “resolve the issue”

Figure 5.4: Chan’s corpus analysis of “problem/issue”

154
On the other hand, worth noting is Chan’s conclusion that “a PROBLEM should NOT be \textit{shared},” while a problem can be shared. In fact, he reported a concern about the possibility of making incorrect prototypes, especially in terms of the analysis of collocate data. As opposed to concordance sentences that can be directly used, frequency collocate data requires use of students’ analytical skills. Chan suspected that students with less grammar knowledge would probably create incorrect prototypes, thus generating false knowledge. In this respect, he wondered about the efficiency of corpus use in teaching the lower levels of students. A more detailed description of his view on collocate data analysis appears later.

\textbf{Attitudes toward corpus use}

Some participants showed changes in their attitudes to corpus use during the course of the study. Recall that Sung and June were initially negative toward corpus use. However, the students’ attitudes gradually changed from skepticism to positive. Their attitudinal change appears to be mainly due to their increased familiarity with the technology. Sung reported that he did not see the benefit of the corpus at first because he did not know how to use the searching query function very well. As he gained more experience with it, he came to enjoy using the corpus. He said: “At first, I didn’t know how to use it, but as I get familiar with using query, I realized that it is helpful. After I used the query better, I can find what I want, so I become more satisfied.” In fact, he became very proficient, as indicated by data collected on frequency of participants’ corpus use (see Table 5.2). Sung was among the participants who used the corpus extensively.
For June, weekly corpus search assignments helped to familiarize her with corpus techniques. By the end of the first term, she felt more comfortable using the syntax searching function and came to realize the value of the corpus as a linguistic resource. In fact, she appreciated the corpus assignments that gave her more opportunities to practice using it. These experiences suggest that students may need the chance to become exposed to corpus technology in order to take advantage of it. On the other hand, June still did not use the corpus frequently, although she reported that she liked it as she became more familiar with it. Her situation is related to her writing process; she consulted the dictionary for definitions before turning to the corpus.

Chan was initially positive toward corpus use and became more positive during the study. His enthusiasm and passion about corpus analysis were quite noticeable throughout the study. He not only checked the corpus to solve linguistic/writing problems, but he also browsed through the corpus, out of curiosity, to test whether the corpus would provide better expressions, even when he had no problem. He said, “I wanted to use the corpus also because I was curious about how well it will produce the results, whether there are better expressions, as well as in order to check whether mine is correct.” Even during the corpus malfunction in the second quarter, Chan regularly (every three days) visited Web sites to see whether it had resumed functioning.

While Chan had a genuine interest in the corpus, the other participants perceived it more pragmatically as a problem-solving tool. They reported that corpus was not something they liked, but they just used it because they needed it to solve writing problems. It was a necessary inconvenience for them. Sung said, “I use it because it’s
helpful rather than because I like it. I don’t like it itself. I think it will be better if we can write without it.” Nick made the same comment:

For me, [the corpus is] not something I like to spend more time on to find words. I don’t like to spend more time on finding words, but when I need, I go there. I know there is information I want, like using a dictionary. I am not saying it’s not useful. But only when I need [I use it]. (Interview 4, 12/2/03)

Xiaodong’s corpus use was also need-based, particularly for writing assignments. He emphasized he would use the corpus more frequently if he had more writing to do because it was only during writing that problems arose. He said:

I am sure I will use it, but I don’t think I [will] use it so much. I may not use it so many times. Just when I need it, I will check it. And if I just search the result everyday, I don’t think it’s so helpful because everyday you don’t know what is your problem in writing. You are only aware [of your problem] when you’re writing. I think if I need to write so much, the more I write, the more I need it. (Interview 5, 3/18/04)

Students’ need-based attitude toward corpus use is a revealing finding. It suggests that technology use should be related to the actual needs of the users. If the technology is not geared toward the students’ immediate needs, it may not capture their interest despite its value in improving their writing.

On the other hand, as Xiaodong insisted, the participants’ frequency of corpus use (see Table 5.2) seemed to be related to the amount of writing tasks they had to complete. Table 5.3 shows each participant’s number of writing assignments for each quarter. All of the participants had at least three writing assignments in the writing class in the first quarter. On the other hand, the amount of writing tasks in the second quarter, when they were no longer in a writing course, varied considerably among them. Table 5.3 shows
that the three science-major participants, June, Nick, and Xiaodong, had no major writing tasks, although they sometimes wrote short lab reports. Of course these three participants used the corpus infrequently. Nick and Xiaodong were initially positive toward corpus use. In fact, Nick was the most positive due to his one-year corpus use experiences. Despite their positive attitudes toward the corpus, however, they actually used corpus infrequently during the study time frame.

<table>
<thead>
<tr>
<th>Name</th>
<th>1st quarter</th>
<th>2nd quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wol</td>
<td>3 writing class assignments</td>
<td>1 case study paper, weekly journal, 2 interview papers, 1 reflection paper</td>
</tr>
<tr>
<td></td>
<td>6 short papers for content courses</td>
<td></td>
</tr>
<tr>
<td>Chan</td>
<td>3 writing class assignments</td>
<td>1 long research proposal, 5-6 page</td>
</tr>
<tr>
<td></td>
<td>1 major paper for content course</td>
<td>long weekly memo</td>
</tr>
<tr>
<td>June</td>
<td>3 writing class assignments</td>
<td>Lab report (mainly mathematical formulae)</td>
</tr>
<tr>
<td></td>
<td>Weekly journal for content course</td>
<td></td>
</tr>
<tr>
<td>Sung</td>
<td>3 writing class assignments</td>
<td>2 research papers</td>
</tr>
<tr>
<td></td>
<td>2 papers for content courses</td>
<td></td>
</tr>
<tr>
<td>Nick</td>
<td>3 writing class assignments</td>
<td>No textual writing</td>
</tr>
<tr>
<td></td>
<td>2 papers for content courses</td>
<td></td>
</tr>
<tr>
<td>Xiaodong</td>
<td>3 writing class assignments</td>
<td>1 short proposal</td>
</tr>
</tbody>
</table>

Table 5.3: The amount of the participants’ writing assignments

Nick and Xiaodong did not even notice the corpus malfunction in the second quarter, while the others discovered it on their own. The main reason was that they simply did not need to use the corpus, as they had few writing tasks. Nick’s papers generally involved numbers and equations; he rarely produced text in a narrative style. In an interview during the second quarter, he said:

I didn’t make so much composition. This is the result. If I don’t have to do composition, I don’t need corpora. I use corpora for papers because for the paper I
need composition, but with my classes, numbers, I don’t need corpora to develop numbers for equations. (Interview 5, 3/10/04)

In addition to having fewer writing tasks, Nick was as a very confident writer whose increased confidence in English writing required less use of the corpus:

I don’t know, because, you know, when you learn more, you start feeling more confident in English, maybe you don’t need to use much corpora. This is one could be the reason to use less corpora. … This is the reason, but when I need to write composition, I have to go there and check my several points. (Interview 5, 3/10/04)

While Nick’s increased writing confidence must have stemmed from other factors, such as more writing experiences and increased content knowledge, corpus use may also have contributed to the increase. If so, corpus use led, ironically, to its own decline. Xiaodong’s comments presented another explanation. According to him:

Although it is not so many times, but some times I need it, so I will use the corpus when I write. It doesn’t mean I do not need it. Just do not need it so much because there are many fixed styles and papers in our field, and we can get such kind of styles and such kind of expressions from other papers. (Interview 5, 3/18/04)

In other words, other sources specific to their academic discipline may have been more useful than the corpus.

**Prototype construction: Participants’ attitudes**

The participants used the corpus as a resource for quickly checking their writing. While they meticulously analyzed the data to construct prototypes for class assignments, they did not engage in prototype construction when they used the corpus for their own purposes. They just checked the search results and solved their immediate problems.
Also, they did not try to remember the results or word usages. Often they searched for the same words again when the same problems recurred. However, this does not necessarily mean that corpus use produced no learning effect. Although the participants did not make an effort to remember the results, learning seemed to take place in the process of consulting searches and solving problems. For example, Nick commented, “you sometimes memorize some phrase[s], not because you want to memorize [it] in your mind, [but because] you [just] remind it.” In other words, identifying problems and finding solutions in the corpus were already part of the learning process. The corpus helped them recognize a problem area, or that some kind of structure is required, even though they did not recall the details of the particular phrase. In this case the students were learning about internalizing linguistic structure/patterns although they forgot the exact details produced by specific searches.

The participants did not feel that prototype construction was an easy task, so they lacked confidence in the activity. In particular, they were not sure about how to use the collocate words that were not shown in the concordance sentences. Chan reported difficulty in dealing with the collocate output, especially finding the sequence between the target word and the collocate word. He was concerned about the possibility of making incorrect prototypes. For example, he said, when the collocate output shows “in” for the target word “both,” but the concordance output does not present how they are actually used in the sentences, he might make incorrect assumptions about those words (e.g., “both in” or “in both”?). He said:

I can still construct prototypes anyway, but I don’t think it’s an easy process. … Although, overall, it’s very easy, if there are one or two very difficult questions, I
don’t feel it’s easy. It’s the same in prototype construction. It’s not so difficult, but because of the possibility of making mistakes, I don’t feel confident in doing it. (Interview 2, 11/10/03)

June also experienced difficulty in putting the collocate words in correct order with the target words. She expressed a high level of uncertainty about proper use of the sequences. She said:

I don’t feel confident in the grammatical aspect. Among the collocate words, I am not sure which words can be put before and which words can go after the target word. I want to make things clear, but sometimes it is not easy to do it. (Interview 4, 12/22/03)

In this regard, Sung believed that the corpus would be helpful only for those who are proficient at grammar and who write extensively. This is an important point for writing teachers using corpora. The technology may need to be adapted to the language skill/level of the user.

Since the participants had difficulty dealing with collocate data, it comes as no surprise to see that most of them mainly used concordances for writing assignments outside the class. The survey, interview, and think-aloud data revealed that they preferred to use concordance data over collocate data. At the beginning of their corpus use, they used the significance data as well, but this became infrequent as time went on. One of the main reasons for them was that it took more time to produce the collocate data than the concordances. Another was the difficulty in interpreting the collocate data. In addition, participants “wanted to look at how words are actually used in the sentences” (Chan), which the collocate data didn’t show them. Another factor was that the user-
The friendly sentence format of the concordances was easier to look at compared to the numerical nature of the collocate data. According to Xiaodong:

> The teacher asked us to include statistics data in the analysis. But, I do not like to do such a thing because usually if I use [the corpus], I just read it without seeing the statistics data because you can see common usage [in the concordances]. For example, there are 10 sentence[s], and I’ve found the same pattern there. So I don’t need to find, to analyze [the data]. I just focus on the parts [where] I am interested. (Interview 4, 12/11/03)

In other words, the participants liked to use the concordances because they did not require data analysis. They preferred to look at sentences as examples of what they wanted within a short time. This was a practical “learn-by-doing” compared to attempting to uncover the underlying principles of the linguistic structure.

From another perspective, the participants’ difficulty in treating collocate data may have been closely related to their striving for “the right answer.” The teacher acknowledge that it is often difficult to make sense out of the corpus results, especially the collocate data. Therefore, he emphasized that students did not need to make neat prototypes and that the main focus should be on the most common usage patterns of words. Despite the teacher’s good intentions, the participants tried to construct “the correct prototypes,” especially for assignment analyses sent to the teacher. In other words, students felt the pressure of submitting the “right” answers, although the teacher promoted corpus use for their own benefit and tried to create a collaborative lexicon. Sung’s comment below confirmed that he looked for “the right formula.” He felt anxious about being judged by the teacher about his analysis. He said:
I didn’t like to submit the corpus assignments. I felt disappointed when my analysis was not put on the class handout. I found that my classmates’ analyses were on the handout, but mine was not. So I realized I didn’t get it right and then I came to dislike using the corpus more. (Interview 4, 2/5/04)

This attitude may have stemmed from the students’ native cultures, where there was an emphasis on finding a single answer among multiple possibilities.

**Prototype construction: A think-aloud activity**

Think-aloud protocols were conducted in order to examine the participants’ processes of corpus data analysis. The think-aloud task was challenging. Despite being given a short practice session before the actual task, some of them did not become very adept at performing the protocols. Some were unable to explicitly describe their thinking process while engaging in another task. In addition, working with corpus data on paper may be different from normal search and analysis online. In spite of those limitations, however, the think-aloud task enabled me to observe the students’ first-hand corpus analysis behaviors that otherwise might have been impossible to be obtained.

The think-aloud task was implemented during the third interview session for all the participants but Chan, who did it in the fourth interview. By that time, they were accustomed to using and analyzing corpus data through prior class assignments. Corpora are believed to be effective in discerning subtle nuances between similar words. “Use (as a noun)” and “usage” were chosen because some participants used them interchangeably in their papers. In the think-aloud task, the participants were given concordance and collocate data of the two words on paper and asked to think aloud while constructing
prototypes of them to differentiate usage patterns. They were given as much as time they needed to finish the task. They averaged 20 minutes each.

The protocols showed that the participants mainly consulted the concordances rather than the collocate data. Only Wol looked at both data sources somewhat equally, as she normally did; she looked at the frequency data first while trying to make sense of it, and then checked sample sentences for confirmation. She checked the concordances to see whether collocate words with high frequency appeared in the sentences. She did some general guessing about the word before she started analyzing the data. For example, “since it is a noun, verbs or adjectives will be followed by the word.” Also, she resorted to her background knowledge while continuing to check whether she heard about the phrases before. She often commented, “it doesn’t sound familiar.” It was observed that she did not hesitate when encountering unexpected data. She made quick decisions based on her background knowledge. She found that “use” still has the characteristics of a verb and acts like a verb, but “usage” is more like an actual noun and combined with another noun to make noun phrases, such as “energy usage” and “water usage.” She concluded that “usage” is a fixed derivative noun from the verb “use.” In so doing, she made a comparison with a corresponding Korean linguistic items to reach a conclusion.

The participant who spent the most time on the analysis was Chan. Accordingly, he produced the most detailed analysis. In a previous interview, he reported that he usually relied on concordance data, assuming that concordance sentences were probably arranged in accordance with their frequency. In the think-aloud task, however, he looked at both data frequently in an attempt to achieve more accuracy about searched items. He looked at the significance data and concordances alternatively, but he still could not
differentiate differences between the two words. Then he looked at the use of prepositions, but he found no specific differences. Noticing that “usage” is combined with another noun, unlike “use,” he then counted the frequency of the form of “adjective + usage” or “noun + usage.” As a result, he found 10 adjective phrases and 6 noun phrases. Then he focused on adjectives afterwards. The following is extracted from his analysis:

The modifiers are different for the two words [use and usage]. Both of them are modified by adjectives, but the adjectives seem to have different characteristics. That is, as “use” mainly has a verb trait, it usually comes with the adjectives that are more like quantifiers, such as “great, little, no, only” or “good, heavy, possible,” which all represent quantity. Other types of adjectives are also seen, like “refused use, intended use, tortured use,” but those are all derived from verbs, not normal adjectives. It means “use” still has the verb trait. In another sense, those adjectives used with “use” can also be seen as adverbs. In contrast, “usage” is commonly combined with “typical adjectives” or “informative adjectives,” which carry meanings, such as, “common, formal, actual, modern” usage. Also, “use” has a wide variety of words to collocate, but “usage” has more limited words to be used together. (Interview 2, 11/10/03)

This example shows that Chan provided a rich interpretation from the detailed analysis. In fact, he was found to have a genuine interest in analyzing the data during the think aloud. After his analysis, he still had some questions about the different uses of prepositions with the two unsolved words. He added that he wanted to search the words again by using different search query items, such as “use of” and “usage of.”

While Chan focused on finding overall differences and generalizing about the patterns, June focused on individual collocate words and tried to interpret every single one. In the meantime, when she ran into an unexpected collocate, she hesitated and did not know what to do. She kept checking with me to see whether her analysis was correct.
In general, she had a hard time doing the task and lacked confidence in her analysis. She stopped the task before getting to the point of generalizing her findings.

In contrast, Sung was very confident in the task, and he followed his own instincts. He interpreted the collocate data along the list of the words and tried to make sense of the words. He was confident about generalizing from his observations within a short time. In fact, he did not go through many words; he just stopped at the tenth word of the frequency data. He concluded that “usage” is commonly placed after adjectives, while “use” comes with prepositions. Also, he found that the two words tend to combine with different kinds of nouns; for example, “usage” often comes with “term, word, or notes.” He also commented that “use” commonly follows “make.” As to whether “make use” was a new phrase to him, he responded that he was aware that “make” and “use” often go together, but this experience taught him that “make usage” or “common use (not common usage)” may sound awkward. It appears that while he relied heavily on his background experience, the corpus analysis expanded, and added new facts to, his repertoire of knowledge.

Xiaodong also relied heavily on his background knowledge. In fact, he did not refer much to the data. Before he started his analysis, he commented, “I remember without seeing the results. I can write down some normal use, for example, ‘make use of,’ ‘use for something.’” He mostly looked at concordance sentences and he seldom checked the frequency data. After he examined “use” and “usage” sentences separately, he said:

There are many normal “make use of,” also “word of use” expressions. Sometimes “use” can be substituted for “usage.” But “usage” cannot replace
“use” most of the time, but “use” can be used like a verb. “Use” is mostly “use of,” but “usage in,” so, “in” is specific for “usage.” … And there are many modifiers for “usage,” but for the “use,” um.. just qualifier, no modifiers, like “high use.” (Interview 3, 11/13/03)

It was surprising to see that Xiaodong did not spend much time on the analysis. He did identify various linguistic aspects of the words, and he was found to be a skillful user of the corpus. This was even more striking because he did not use the corpus frequently, in which case it seemed safe to assume that he was not used to performing corpus analysis.

As for Nick, he did not adopt an efficient or refined strategy in his analysis. His think-aloud observation suggested that he was an unskillful corpus user, an unexpected finding given the length of his exposure to it and the enthusiasm he expressed in the interviews. His examination of the data did not come to any specific conclusion because he was not analytic or systematic in his approach. For example, he found that “‘use’ is for kind of general thing, but ‘usage’ is for not so general.” Basically, he treated the data separately, rather than combining it.

The participants’ think alouds reflected their levels of effectiveness in using the data, their analytic strategies in examining the data, and their familiarity with prototype construction. While the participants showed differences in the depth of their analyses and interpretations, they all readily recognized the pattern “make use of,” the popular chunk phrase taught in English courses. That is, they recalled and confirmed their prior knowledge with little doubt about it. This, then, raises important questions as to the role of corpora in language learning. Does the corpus approach correct prior knowledge, or does it just confirm and expand upon it? More importantly, it is questionable whether
participants without prior knowledge would find the pattern themselves from their own corpus analysis. This also leads to a question about the efficiency of the teaching method: whether to present the students with a ready-made chunk or to let them find it through an independent search.

Those questions are later discussed in this dissertation. For the moment, the think-aloud protocols showed the participants’ capability as sophisticated corpus analysts. Their processes conformed to Johns’ (1991) description of the basic procedure of concordance-based learning as “identify-classify-generalize.” The participants in this study identified unique features, categorized them, and finally constructed common usage patterns from corpus data analysis.

**Alternative uses of dictionary and corpus**

The participants used the dictionary and corpus for different purposes: dictionary for definitions, corpus for usages. Many of them used the dictionary as the first reference when learning definitions. They used the corpus as the second resource, mostly for finding word usage.

Initially, many participants preferred the dictionary to the corpus because of its convenience, including its portability and greater speed of use. For example, Wol believed that “a dictionary can take care of many of [my] problems.” Later, however, they began to prefer the corpus. Xiaodong commented, “corpus is better than [a] dictionary because sentences come from common usage and there are many sentences. [A] dictionary has only one or two samples.” Chan relied on a dictionary for word usage
until he got used to the corpus. Then he concluded that “a dictionary can’t do much about usages and connections between words.” Sung agreed:

Now problems are more about usage rather than meanings, so I think about using the corpus first rather than dictionary. I am thinking about the corpus first for word usage. (Interview 4, 2/5/04)

Nick clearly differentiated dictionary use from corpus use. He did not use them interchangeably. That is, he did not use the corpus for meanings, and neither did he look up usages in a dictionary. He commented:

If I need some meaning, I can go to the dictionary to find the meaning. I used the corpus in the past to find the meaning, but not so much. I feel more comfortable with dictionary. (Interview 4, 12/2/03)

On the other hand, Nick added:

I know almost of meanings, but the problem is becoming more connection between phrases. Dictionary is not helpful for that. Maybe dictionary is helpful, but corpus is easier, more comfortable. (Interview 4, 12/2/03)

As advanced learners, the participants did not need the dictionary that much because they had few vocabulary problems. Therefore, they used the dictionary less often as the study proceeded. In fact, some of them seldom used a dictionary as the study reached its conclusion. By that time, they needed more usage knowledge rather than meanings of words. As such, their frequency of corpus use increased.

While the dictionary and corpus served different purposes, the participants also used them complementarily. In fact, the corpus became a second step resource after the dictionary, rather than a completely separate resource. Students first consulted the
dictionary for meanings and basic usages, and they then checked the corpus to deepen their usage knowledge. Nick said, “if something in the dictionary is not so clear for me, then I can find on the corpora.” For Chan, corpora were used mainly for important decisions that might have an effect on language learning/writing. For example, he consulted the corpus to differentiate the usages between “on the contrary” and “to the contrary,” phrases that can be used often in writing, but dictionaries present few examples of their usages. June used a bilingual dictionary to a great extent in order to come up with new English words to use. She could not use the corpus at first. Since finding words to write a sentence was her first priority, she had to use the dictionary first. According to her:

The corpus is not so helpful for me so far. The thing is that I don’t know what words to use at the first place. In my writing I get ideas in Korean first and then translate them into English. I usually go to an Internet bilingual dictionary which gives me corresponding English words. Then I can check the English words in the corpus to see what their usage patterns are. So, at the beginning, the corpus is not helpful at all. After I make sentences, I can use it to make changes to them, but usually I don’t have time to get to that point either. (Interview 2, 11/5/03)

Interestingly, students did not revert to their dictionaries when a technological breakdown made the corpus unavailable to them. By that time, corpus use had taught them that the dictionary could not solve problems that were mostly related to connections between words, rather than their meanings. In a roundabout way, corpus use taught students the limitations of dictionary use. For example, when the corpus was unavailable, Sung left unchecked certain parts of his compositions. Chan also passed over problems without even looking them up in the dictionary after the corpus breakdown. He said, “I
do not even try [to use] a dictionary. I think, although I looked up a dictionary, it may not solve the problem.”

**Beneficial areas and the roles of corpus use**

Some participants used the corpus while initially composing their writing assignments, while the others used it only for revising. Interestingly, the more frequent corpus users, Wol, Chan, and Sung, used it for composing. The less frequent users’ main purpose was in revision. June, in the latter category, reported that a key factor was its availability at the time of writing. However, she preferred using a dictionary for initial composing while reverting to the corpus for revision. Nick delayed his corpus use until revision because he wrote his first draft by hand, and then typed it into a computer, at which time corpus was available. He often marked problematic words in his handwritten draft, and then checked them in the corpus while revising.

Among those employing corpora while composing, Wol usually kept the corpus Web site on the corner of the computer screen while writing. Chan also checked against the corpus in writing whenever possible. Sung, who changed from an unfavorable to a favorable user of the corpus, was a careful writer. He used the corpus during writing and reported that stopping to check the corpus while writing did not seem to hinder the flow of his writing; Sung reported it was a familiar pattern from his reading experience, i.e., stopping to check a dictionary during reading. Chan responded that consulting corpora interrupted his train of thought somewhat, but the effect was negligible, maybe “only 10 out of 100 points.” Chan believed that interruptions were worth the advantage of making his revisions much easier:
If I don’t check the corpus in the middle of writing, when I revise I have to go through many stages of revision. And even though I revise many times, it’s hard to catch the mistakes later because although I plan to look them up, I often forget or skip them later. So, checking while writing eases my revising process. (Interview 2, 11/10/03)

With respect to the benefits of corpus use, most participants agreed that it helped in writing, but was not useful for reading or for learning grammar. They did not equate the corpus searches with reading activity because they did not provide an extended amount of text beyond sentence-based concordances. In the same light, June and Nick commented that even for writing, the corpus was helpful only for sentence-level writing, not at discourse or organizational levels. Nick said, “corpus is good for sentence level, but [the] class is good for content.” His comment is insightful. In fact, it reflects the teacher’s pedagogical model for the class. Corpus is not an end in itself, but is only a tool in a holistic teaching process. This model employs corpora for sentence-level writing problems, which students solve for themselves, as encouraged by the teacher. The teacher uses other materials and other activities to teach organizational and rhetorical aspects of writing. As such, the combination of the corpus use plus classroom instruction contributes to students’ successful writing experiences. In this way, the participants viewed the corpus as a valuable resource for addressing their linguistic problems. In fact, the participants did not express serious dissatisfaction with the sentence-focused corpus orientation; they accepted it as the nature of corpora.

When corpora are viewed as a linguistic resource, one important question that arises is the precise role they play in language learning. This study’s data strongly
indicate that the participants used the corpus to reinforce and confirm their prior knowledge, rather than generating new knowledge.

The participants themselves asked profound questions regarding whether new learning would take place by corpus use. Their comments were revealing. Wol suspected that corpus use did not create new knowledge because she only used it to double-check or review whether her current knowledge was correct. She only used the corpus data selectively based on her prior knowledge. That is, if the data was inconsistent with her current knowledge, she just ignored it and did not expand her range of linguistic performance. Her learning curve remained static. She said:

When I am looking for, say, a phrasal verb, I don’t do guessing in advance. From my past rote learning experiences, I always try to recall what I learned. Even when I look up the corpus, I try to recite my memory from the rote learning, rather than guessing based on the data. So, for me, even guessing or checking just means checking whether my memory is correct. (Interview 5, 3/11/04)

Chan asked the same question: “Which comes first? Do I analyze the data from the scratch or do I interpret the data just based on my prior knowledge?” He continued:

When I am analyzing corpus data, especially the collocate output, I often ask myself the question whether I think my prototype is correct based on the data or on what I’ve learned. I seem to use and interpret the data based on my prior knowledge. Therefore, I doubt about whether I can interpret the data and construct prototypes if I don’t have any grammar or word knowledge. (Interview 2, 11/10/03)

Sung used the corpus to “clarify” (not newly learn) the exact meanings of words. Nick also said that corpus use reinforced and improved his grammar knowledge, but it did not help him to acquire grammar. According to him:
[The corpus] is not the tool to learn grammar. It’s helpful, but not the tool for grammar. [By] using the corpus, you cannot learn grammar. It helps grammar, but you cannot learn grammar from the corpus. (Interview 4, 12/2/03)

Certainly, people refer to their prior knowledge to acquire new knowledge, and corpus analysis requires a certain level of prior grammar and word knowledge. As noted earlier, an interesting question is whether corpus use only confirms, reinforces, or even serves to review the student’s prior knowledge, or does it extend the process to new learning? This is not only an important question for this study but has broad ranging implications regarding the use of technology in education in general. It poses severe challenges to a techno-deterministic learning paradigm, particularly as universities increasingly offer online degrees.

One clear finding from the study was that most of the participants did not use the corpus for learning new material. This point becomes more evident if we look at their corpus use behaviors. They rarely used the corpus for unknown words or phrases; rather, they used it for checking or extending their current knowledge about words. This might not seem surprising given their advanced level as L2 users. On the other hand, many of them reported that they checked the dictionary for definitions of unfamiliar words in order to obtain new information about meanings and usages. They only used the corpus to solve their linguistic problems. They viewed the corpus as a linguistic resource like a dictionary, but while they perceived a dictionary as a tool for new learning, they viewed the corpus as a secondary source, a linguistic checker.

The question of whether learning takes place through corpus use also raises the issue of how we define “learning.” Does learning refer only to acquiring unknown facts
and adding new knowledge, or does it also include checking uncertain knowledge, reinforcing prior knowledge, and finally internalizing it as one’s own knowledge? While this issue may require further discussion, it is not an understatement to say that language learning is not a one-step event. Rather, recitation and reconfirmation are integral parts of the whole language learning process, as reflected in the commonly recited adage of foreign language teachers: “use it (language) or lose it.”

In addition to providing actual textual help in writing, the corpus helped the participants gain a psychological advantage in their language learning. Once they checked assignments against the corpus, the students felt more confident about their L2 writing. Chan said, “if I look up the corpus, I feel confident, regardless of whether it’s correct or not, because once I checked it, I feel it’s right and I feel comfortable.” Xiaodong also said: “I have confidence because you get common usage they are always used this way. I am sure I am following the common usage.” According to Nick: “Because you know you have 40 sentences [from which] you can pick the best phrase for your phrases, you feel confident because it’s not mine and I know it’s right.” That is, if they consulted the corpus, they felt confident about using the common usage correctly, and this enhanced their confidence in their writing. On the other hand, as noted earlier, Nick added that more confidence in English writing would, ironically, lessen his need to use corpora.

5.1.2 Corpus use, problems/difficulties

In addition to the participants’ overall corpus use patterns and their perceptions of its benefits, this study also examined students’ difficulties with corpus use. This is an
important aspect of the study because it was essential to avoid the pitfalls of adopting a
techno-deterministic viewpoint, i.e., that all technology bestows unlimited benefits to its
users. For a study of educational technology to be truly useful, we must necessarily
include its shortcomings and then weigh the costs and benefits. This is also necessary in
order to create a fuller and more accurate portrayal of corpus use. This study covers a
wide range of potential problem areas, from technological aspects to corpus data
characteristics to the effects of student’s level of English proficiency on learning.

**Technological difficulties**

The participants agreed that corpus availability/portability was a decisive factor in
their corpus use. This factor includes not only whether they had a computer or Internet
connection, but also, whether the corpus was available at the time of writing. Wol said
portability was the most important problem in corpus use; when she lacked an Internet
connection when writing, she reverted to a portable electronic dictionary. For Chan, the
Internet was unavailable when he wrote at home, so he could not check the corpus
although he wanted to do so. Chan said:

> Sometimes I write papers at home. For example, I wrote a draft of the writing
class assignments at home and took it to the teacher. He asked me why I used
some expressions which seemed to be uncommon. It was because I wrote at
home. If I had written it at school, I would have checked them in the corpus. But,
when writing at home, I don’t check it although I knew some expressions are not
quite right. (Interview 2, 11/10/03)

June also complained about the inconvenience of being required to be connected
to the Internet when writing, which was sometimes impossible. She preferred a portable
dictionary.
Participants also had trouble with slow Internet connections. Wol’s home Internet speed was so slow that corpus searches took too much time. The speed of the Internet was also too slow for Sung. He also encountered a computer breakdown that disrupted his writing at the beginning of the first term. He could not use the corpus until his computer was fixed one month later. Another problem was that Sung could not get what he wanted from his corpus searches. He provided these reasons: 1) the corpus did not have the desired information, 2) he asked the kinds of questions that native speakers would never raise, so the native speaker data could not answer all his questions, or 3) he did not know how to search for the data. Sung was not the only one who showed concern about difficulties with the searching techniques. Wol and June also complained about difficulties in using the technology. Wol did not have any serious difficulty, but she was still not familiar enough with the searching queries to conduct more sophisticated searches. June, who became somewhat accustomed to the technique later, still felt the need for learning more about the technology in order to use it more for her own specific purposes.

Ironically, overall, the participants who used the corpus more frequently talked more about the difficulty or problems in corpus use, while the infrequent users provided few complaints. Wol, Chan, and Sung, who were frequent corpus users, noted difficulty with the corpus availability, Internet speed, and searching technique, but Nick and Xiaodong, who used the corpus least, reported no major problems regarding the technology. Nick explained that Internet connection was no problem for him. Xiaodong did not have any trouble with Internet speed because he mainly used the corpus at school, where the connection to the school server was a great advantage. That frequent corpus
users encountered more problems than infrequent users is, however, not a surprising finding, since greater use exposes one to more opportunities to experience difficulties. This situation is problematic because we do not know 1) if infrequent users are discouraged by problems and hence avoid the technology altogether, or 2) infrequent users would indeed have more complaints if they had used it more. On the other hand, these patterns may also show that the difficulties were not serious enough to hinder the students’ corpus use.

**Data presentation**

Some participants complained about the way corpus data results were presented. For Chan, it was inconvenient that concordance data and collocate data are not provided on one page, and that each requires a separate search. He believed that if the two types of data can be seen at the same time, it would shorten the search process and make it more efficient. Sung made the same suggestion:

The corpus is inconvenient because I need to go to two different Web pages to get the full information. I think it will be great if they can show the results at the same time. For example, they can produce the data at one page, or with a click on a button, it can bring a pop-up window to show the frequency. (Interview 5, 3/9/04)

Chan also pointed out that corpus data presentation can be much too mechanical in nature and lacks the necessary context for solving a specific problem. According to him:

The corpus is too much mechanical. It mechanically produces a bunch of data which is sometimes irrelevant. It doesn’t give us interpretation, although it gives us many helpful examples. I found the frequency data is sometimes incorrect. Some collocates, which are not commonly used, show a higher frequency in the
data, and there are expressions that should not be used that way. So, I am worried that a person with low grammar knowledge may come up with wrong prototypes from the data. (Interview 1, 10/28/03)

With respect to the data presentation in cut-off sentences, participants expressed similar attitudes. Most of them mainly concentrated on the middle part of the sentences rather than the whole string, which showed their focused or narrow use of the corpus data. Wol did not have any trouble with cut-off sentences because, generally, she checked only the part she was looking for rather than trying to read the whole string. It was the same for Sung; he also had little trouble with it. Xiaodong just focused on the part in which he was interested in order to quickly find the answers to his questions. But he still believed that whole sentences would be better because “sometimes, the word, the meanings of the word are not clear without the whole sentence,” and “the language environment is important to decide which kind of combination you will use.” He continued that he would look at the whole sentence if he had time. For Chan, the cut-off sentence presentation was not irritating generally, but it was unsatisfactory when he needed a longer textual context for clarification.

Only June reported that she checked the whole line of the sentence in her corpus use. She said:

I check the whole sentence for studying English. I believe it’s helpful for learning English. Just living in America does not guarantee better English skills and I’ve found that I don’t study English as much as I did in Korea. I need to study English, and that can be one way. (Interview 5, 3/10/04)

It appears that the way the students’ used the corpus diverged depending on their purpose for using it. That is, while the other participants perceived the corpus as a
reference for checking their questions quickly, June seemed to also view it as an English language learning tool.

**Time-consuming data search and analysis**

A few participants (Wol, Chan, and June) complained that a serious problem in corpus use was the time-consuming nature of data analysis, but the data indicated that the cause of “time consuming” aspects varied. These ranged from technological factors to student language proficiency. The participants referred to the time consuming nature of both data searches and analyses. First they complained about the actual time required for getting access to the corpus, searching for the data, and producing the results. In particular, they had to browse for two different types of data due to the lack of a user-friendly interface (i.e., consumed concordance and collocate data). Once they had access, they consumed time performing data analysis. The data had to be examined, categorized, and translated into prototype strings. In so doing, language proficiency plays an important role in making sense of the data. For example, data analysis took June so much time that she wanted somebody to do it for her and just present the results to her. In other words, what is required in corpus use is not only efficiency, but also language proficiency in the search for and analysis of the data.

**Time availability**

Time was a crucial factor in the students’ attitudes toward corpus use. Throughout the study, Xiaodong stressed that he did not have enough time to use the corpus. He highly praised the corpus, but he had no time to use it. He made an insightful analogy that “I believe it’s very helpful, but I don’t have time to use it. It’s something
like, you have many good videos, movies, but you don’t have time to see them.” He rejoined, “you know, I have not finished even assignment, so how can I use [the corpus]?” In fact, improving writing skills was not his first priority. He said:

If I only want to [improve] my writing, maybe I will use [the corpus] more, because it is very helpful, and you can get many types of beautiful sentences from it. Now I don’t think I need it. I don’t have time to improve my writing skills. Um, so far it’s not so important as other stuff, for example, if you have a class, you need to pass. If you fail, it is terrible, you know. That’s more important than writing. (Interview 5, 3/18/04)

That is, Xiaodong felt less urgency about writing improvement, which could be delayed, so he could focus on his content course assignments and coursework. His decision may have been partly due to his science-based major that requires less writing, so he did not consider writing to be an integral part of successful completion of his doctoral courses. On the other hand, he added that he would use the corpus for formal papers in the future. In other words, while time constraints did not allow him to use the corpus at that point, he was aware of its benefits as a future reference.

June also pointed out the problem of time availability. She believed that if time allowed, corpora would be useful for studying English. But given time limitation, the dictionary was more efficient for her purposes.

Interestingly enough, while all participants were busy graduate students, some still spared time to use the corpus. Thus, in another sense, the time issue could also be related to the students’ need of and satisfaction with the corpus. One might assume that students who are highly satisfied with corpora would be willing to devote time to corpus use.
English too general for disciplinary needs

Sung explicitly complained about the general nature of the corpus and expressed the need for a specialized corpus in his field. He reported that the corpus was not helpful for writing about history, his discipline. On the other hand, he came up with his own solution by adapting the corpus approach to his needs. He kept his own collection of expressions, words, and phrases taken from reading articles and books related to his major. In effect, he constructed his own small, specialized corpus for his field. In fact, most of the participants (Wol, Chan, June, and Xiaodong) kept a notebook in which they recorded phrases and sentences from their reading. In other words, the learners were creative adapters of the corpus approach to their studies in a broader sense.

While Sung desired a corpus particular to his academic field, Wol and June pointed out the need for a corpus for the broad genre of academic writing. Wol indicated in the survey that the corpus did not help her with academic writing, because the results often included informal speech data. She believed it would be better to read textbooks and journal articles to acquire an academic writing style.

Although June did not specifically mention the need for a specialized corpus, her corpus search examples illustrated her frustration toward the general corpus. She actively searched for academic words, such as “a priori,” but the corpus did not produce much relevant data. While her survey response indicated that she did not need a specific corpus, she expressed a desire for one in a later interview:

I hope I can select only science materials or a category that is relevant to my field, say, not from newspapers. If I can search through only those materials, then it would be faster. Although it is not helpful for learning general English, but it would be helpful for writing my papers. (Interview 5, 3/10/04)
On the other hand, Chan, Nick, and Xiaodong were not dissatisfied with the general corpus. Nick’s corpus use experiences taught him that it was not useful for technical words, but only for common words. So he already knew it was not useful for words used in his discipline or academic writing in general. He said, “you can’t find technical word on the corpus. I don’t even try because the corpus shows the most common words.” At the same time, he appreciated the fact that the general corpus helped him with prepositional usages and common usage patterns. Xiaodong overcame the corpus limitations by using a specialized dictionary for technical words. Because the dictionary helped him to deal with his needs for his field, he did not demand such information from the corpus. He believed he did not need to use corpus anyway, because his major did not require a variety of writing styles. He indicated he could learn the styles in his field by consulting academic papers.

As seen, the students adjusted their expectations based on their experiences with the corpus. Likewise, they learned its limitations and capabilities and adjusted their uses accordingly. What we see here, then, is a creative adaptation that extends beyond the bounds of the corpus developers. It reiterates a continual theme running throughout this study. Students are not passive receptors of knowledge, but rather they engage material with cognitive skills that shape the learning process. Understanding this process has wide-ranging ramifications in any sort of learning experience beyond the bounds of corpus research. A fertile ground of research could extract these general principles and fruitfully apply them to other pedagogical areas. Great strides could be made in individualized learning styles by following up on this procedure.
The teacher, on the other hand, was not a passive bystander while the students grappled with their disciplinary needs. The teacher actively catered to the different needs of the students from various academic backgrounds. For example, in one class assignment, the students created a style manual by referring to journal articles in their own fields. Students were asked to find five articles from representative journals and to describe the unique features in each section of the article, such as the introduction, methods, results, and discussion portions. By doing so, they had to cite commonly used phrases. This class assignment itself can be seen as a type of corpus construction in the miniature. It offers complimentary and alternative material to the general English corpus, showing discourse-level characteristics of disciplinary writing beyond the sentence-level text. Given the various disciplinary backgrounds of the students, this was a sound attempt to cater to the individual needs of each student in the class. Whether this approach should be expanded and developed is a worthy topic for future research.

**Proficiency level and independent learning**

As reported earlier, June had a great deal of difficulty in corpus use. She had to consult the dictionary first before using the corpus. She was deeply concerned about language learning, so she approached the corpus as a language learning tool. The other participants saw it mainly as a more practical resource. However, the corpus fell far short of serving June’s need for language learning because it was too time-consuming.

However, her search logs and interviews indicated difficulties in casting the appropriate questions to the corpus. The following examples are extracted from her corpus searches:
1) – “journal article”:
I did not find what I want. Maybe I don’t know exactly how to use it, but I
couldn’t find it. I looked for “journal article.” Generally, “journal” means a book,
or an issue of the collection, right? But I am not sure whether I can just say
“journal” or “article in the journal” and when it’s used as a proper noun, for
example, “wall street journal,” I don’t know whether I need to put “the.” Some
technical journals are called “the journal of…”
2) – “direct quote”:
I wanted to know how we can express “direct quote” when we speak. So I
searched it, but found only two results.
3) – “explicit/implicit”:
I wanted to know the exact meanings of “explicit/implicit.” But the first word that
I got was “sexual.” Does the word have a different meaning? Since I saw the word
“explicit” in a textbook, I thought it is a formal word, but I was surprised to see
the word “sexual” with it.
4) – “a priori”:
What is the exact meaning of “a priori”? Is it similar to “assumption”? I didn’t
get what I want in the corpus. (Search log, 11/11/03; Interview 3, 11/19/03)

The searches indicate that June mainly looked for word definitions in the corpus,
which failed to accomplish her purpose. Her corpus use differed from the others who
focused mainly on finding word usage. In this sense, she did not know what to expect
from the corpus and how to conduct an efficient search. A possible reason is that she was
not used to the searching technique and perhaps needed more training. On the other
hand, her difficulty in corpus use also seemed to be related to her lower level of language
proficiency. During the study, she kept checking with me about whether her search and
analysis were correct. As such, she reported that she wanted someone to help who would
interpret the results and make patterns for her. She said:

Finding a pattern and making a prototype from the corpus results takes me too
much time. I hope somebody can do it for me and just tell me the common usage
patterns, e.g., “make use of.” If I am fully trained to use the corpus, I may be able
to analyze the data, but I can’t do it now. I think it would be the most efficient if
anybody can give me the result patterns. I don’t think I can do it by myself.
(Interview 4, 12/22/03)
June’s comments illustrated that her grammar or English proficiency level was inadequate to analyze the data, and she needed outside help for doing it.

Another possible explanation could be related to her past learning experiences. As she was presumed to have been taught by a traditional transmission model, any proactive, independent learning, or learner autonomy, was not reflected in her corpus use. As a passive learner, she may have wanted to receive ready-made results from others rather than analyze the data herself. This point, then, casts doubt on corpus researchers’ argument that corpus use enhances learner autonomy through an independent problem-solving approach (Johns, 1991; Thurston & Candlin, 1998). What is not clear here is whether corpus use would enhance independent learning, or vice versa, that is, whether only already autonomous learners would like to use the corpus. While this point calls for more discussion, it seems likely that students’ corpus use behaviors and level of satisfaction are also closely related to cultural variables regarding learners’ expectations in the teaching and learning process.

5.1.3 Students’ overall evaluation of corpus use in L2 writing

At the end of the study, the participants were asked to write reflections on their corpus use experiences during the study. Chan and Sung chose to write in Korean, while the others chose English. I translated Chan and Sung’s reflections into English for the purpose of this study. As their reflections validated and repeated the main points made during the interviews, this section is mainly devoted to presenting their own words.

Wol recognized the benefit of corpora for acquiring word usage, while she pointed out the time-consuming problem of getting access to the Internet in general, and
its availability at the time of composition. The following was extracted from her reflection notes:

… The advantages of using the corpus are providing various example sentences that used the targeted words and providing frequencies of the use of the words. Through reviewing the example sentences, I could learn the usage of proper prepositions and verbs. … The difficulty of access and time-consuming process are the disadvantages of using the corpus. I cannot use the corpus every time since the corpus is provided through only web. Therefore, if I cannot access to an Internet, I cannot use the corpus. … (Written reflections, 3/25/04) [original in English]

Chan, in his lengthy reflections, remarked on the advantage of corpus use for identifying proper modifiers and common usage patterns. He found it so useful, he introduced the corpus to his friend who was struggling with English:

… Two weeks ago, I had a chance to talk with my friend who is studying about art critic. I found that he had difficulty in English writing and introduced the corpus to him. What I introduced to him was as follows: “There is an Internet site which stores commonly used expressions that we can search by frequency. It is convenient to check which words are used together, which is helpful for English writing. In my case, I used to be interested in which prepositions come with certain verbs, but actually we can find the information in a dictionary. On the other hand, we cannot easily find proper adjectives and adverbs, for which corpora can do a better job. It can be helpful for art critic or my study that requires delicate nuances for more successful communication, beyond simply transmitting a message.” … I have trouble with the corpus website breakdown, particularly when the dictionary does not have enough usage examples for proper adjectives or adverbs, or when I want to check whether the expression is common after composition. (Written reflections, 3/1/04)

June’s reflections indicated her gradual familiarity with the technique, which was the initial obstacle in her corpus use. She came to appreciate the technology over the course of time, so she even felt “uneasy and irritated” when it malfunctioned later.

According to her:
… At first time when I used CCC [the corpus], it was very inconvenient to me. It didn’t instantly give me the exact result what I expected. The reason maybe was that I wasn’t accustomed to the query syntax. It was also time-consuming procedure to find the prototype of several examples among the corpus concordance sampler. The unfamiliarity with it made me use another on-line dictionary which provides me quick answers. … Fortunately, our instructor let us use CCC at least once a week and make a prototype from the query results. By doing that assignment, I became familiar with using it. I found the difference and uniqueness in CCC in comparison with other dictionary. It shows me the contemporary words in every media and a shade of difference in meaning and usage of words. [Italics added] … Now, I have been trying to make the most use of it. When I fail to access to CCC web site, I feel uneasy and irritated. I am looking forward to use it again sooner or later. (Written reflections, 3/17/04) [original in English]

June appreciated corpus search assignments as an opportunity to become familiar with the technique. Sung disliked the assignments, but he found the value of the corpus in his writing. In a brief note, he reiterated the need for a specialized corpus in his own field:

The corpus is good for learning how to connect words and what are the common usage patterns of words. But sometimes I could not find what I wanted in the data. Also, I was able to understand most of the data, but some cases I could not understand because of search errors, which may need some improvement. In short, I will continue to use the corpus, and hopefully, I hope I can have a specific corpus in history. (Written reflections, 3/5/04)

Nick viewed the corpus approach as an alternative approach to “the traditional methods.” He praised the corpus for improving writing and vocabulary in general and learning verb usages in particular. He said:

… I learned how to use corpora and collocations. From that point I start to explore English, besides the traditional methods, in a different way by searching the most used pattern. … One of the most functions that I used was pattern of the verb with a conjunction. Also, Collins Cobuild was very useful for me, because in comparison with dictionaries, it showed the most used pattern for a word in
desired form such as verb, noun, adjective, etc. … I have confidence in using Collins Cobuild because in my English class we discussed the findings and they were very comprehensive. In fact, with each search I found a new pattern or phrase that improved my writing and vocabulary. (Written reflections, 4/23/04) [original in English]

Nick’s corpus use behaviors deserve elaboration, given that he had the longest period of exposure to corpus among the participants. He showed a very positive attitude toward corpus use since he was involved in the prior study preceding the current research. He did not report any major difficulties in corpus use, saying “no difficulty, everything is great!” However, he did not use the corpus as frequently at a later stage of the study. In fact, his frequency of corpus use was eventually quite low. In contrast to his professed confidence in corpus use, he was not so skillful in prototype construction. In addition, his searches were limited to prepositions, while the others’ searches showed a wider variety of corpus use and expanded to include wider linguistic features. One reason for his infrequent corpus use was that his science-based studies required few writing assignments. He simply did not need to use the corpus, except for required class assignments rather than by choice during the study. Because he had used it for one year already, its novelty had worn off, so he lost interest in exploring it any further. Another reason for his lack of interest was that his increased knowledge about writing and English meant that he needed the corpus less and less as time proceeded.

Xiaodong acknowledged the advantage of corpus use for learning common usage patterns, although he was an infrequent user of the corpus. Notably, he remarked that corpus use would decrease the influence of native language usage on English writing. In
this regard, he argued for introducing the corpus students at an early stage of English learning:

Although I did not use corpus frequently, I think it’s a useful tool for English study and writing. I studied English more than ten years. However, I never found a useful tool that can provide the common usage of words and common expressions used by native English speaker. At the early stage of English studying, it’s very important to be used to the common usage in native English, not the usage based on our own language. It can help us to decrease affection from our own language. That’s really what our non-English-speakers needed. So I think the corpus is very useful for the beginner of English study. Another use of corpus is for our formal writing. There are different using preferences for different content. [Italics added] For example, when we are not sure which kind of usage of words that have similar meanings, should we use, the corpus could provide excellent comparisons and then we can get right information. (Written reflections, 4/9/04) [original in English]

Xiaodong comments in italics show that the corpus helped him with formal writing. Recall here that June also made the comment, “it shows me the contemporary words in every media.” In other words, she also recognized the value of the general corpus. Generally speaking, then, the students saw the general corpus as an advantage rather than a limitation in pursuing English writing competency. The main point, then, is that despite students’ various difficulties in using the corpus, their overall view was favorable with respect to its usefulness.

To summarize, this section reviewed the case study participants’ corpus use patterns and their perceptions of its benefits and difficulties. It is worth noting that the students expressed similar points regarding advantages, i.e., learning common usage patterns and showing subtle nuances and contexts of use. But students’ perceptions of difficulties diverged, ranging from technological issues to corpus characteristics, to time availability. As for problems, the main obstacles were as follows: for Wol, the time
consuming aspect of accessing the Internet and Internet availability at the time of writing; for Chan, slow Internet speed and technology malfunctions; for June, search difficulties with techniques and her own presumably less proficient English skills; for Sung, the too general nature of the English of the corpus; for Nick, no major difficulty, but a lack of relevance because he had few writing assignments; and for Xiaodong, insufficient time to use the corpus. The data demonstrate that although the participants were taking the same level of writing course, they had different issues and problems. These involve technical skills, disciplinary content knowledge, English, and writing proficiency.

Also noteworthy was the relationship between participants’ corpus use frequency and the number of writing assignments they had to complete. The participants’ corpus use decreased when they had fewer writing assignments. In fact, they asked for more writing assignments in order to increase corpus use. Nick and Xiaodong, who did not use the corpus that much in the second quarter, planned to use the corpus more when they expected to have more writing assignments and formal papers in the future.

Now that the students’ interactions with the corpus have been addressed, the next section examines the influences of corpora on students’ understanding of the English language, and their attitudes toward English writing after they adopted the corpus approach.

5.2 Influence of corpus use on students’ understanding of language and L2 writing

During the course of this study, the case-study participants’ perspectives toward language learning and L2 writing changed. Because these changes occurred after their
involvement with the corpus, it presumably was an independent variable altering their perspectives. Their initial perceptions of language and writing experiences discussed in Section 4.1 were compared to the perceptions examined later in the study. In addition, their changed perceptions were explored via interviews at the end of the study. This section presents the findings concerning their understanding of language and language learning acquired through the corpus approach, and then it moves to their later approaches to L2 writing and writing processes. It is necessary, though, to be cautious about placing too much emphasis on the antecedents or determinants of the changes that were observed in the students’ perceptions. Unlike an experimental study that controls other “confounding” variables, this study was conducted in a natural environment. In a natural inquiry like this, it may be misleading to conclude that the students’ perceptions of language and L2 writing changed only due to corpus use experiences. The relatively short length of the study is another limitation. The six-month duration, particularly given the unexpected technology breakdown, which occurred during the last three months, makes it difficult to identify any conclusive evidence of changes in writing practices. On the other hand, certain similarities appear that point in the direction of corpus’ influence. These caveats need to be remembered in the following exploration of any prominent influences of corpus use on participants’ L2 learning/writing experiences.

5.2.1 Changes in the students’ understanding of language and language learning

Increased perceptions of lexico-grammar

With respect to understanding of the language, most participants expressed a combined concept of vocabulary and grammar, or “lexico-grammar” as it is called in
technical terms. Recall that Chan was initially placed at the lexis-bound end of the continuum, with grammar on the one end and lexis on the other (see 4.2: Views on language and language learning). He believed that explicit grammar instruction is unnecessary beyond a certain basic level, while arguing for descriptive grammar. Later, though, he showed an appreciation for grammar. His attitude toward systematic grammar instruction changed. While he held his original views toward descriptive grammar, he recognized the value of teaching systematic grammar. According to Chan:

Learning a language is to learn how the people of that language use the language. Recently, I talked with my brother-in-law who is studying English in a language program. He complained that when he spoke based on the grammar he learned from high school, his teacher told him “we don’t say it.” I think his grammar could be right, but it was just different from the grammar that the teacher uses. I think people tend to believe that the way they speak is grammatically correct. Grammar is being changed and modified to reflect people’s uses of the language, but rules are always late to follow up the trend. The purpose of language learning is to communicate with the people of the language, so we need to learn the ways that are most widely used. Those ways are changing and sometimes not logical. Therefore, in contrast to what I said earlier, we can’t say learning a systematic grammar is meaningless. In other words, we can’t say that the grammar we learned in Korea is useless. (Interview 4, 3/8/04)

However, his increased concern for grammar did not mean that he agreed with the traditional concept of separation of grammar from vocabulary. Rather, he viewed grammar more through the concept of collocation. He believed that what he had learned as grammar was, in fact, closely related to collocation. As he explained:

Basically, what we learned as grammar is all related to collocation. For example, we just learned “make use of” as a chunk, but the fact that it is not “make use in” or “make usage of” is based on collocation. Likewise, sentence structures or common expressions are all related to collocation, but we did not use the concept “collocation.” Instead, we called it as “grammar, accurate usage, or idiomatic expressions.” Nobody taught me about collocation, and moreover, I didn’t have
time to care about it. However, collocation is one of the best ways to find the most appropriate and beautiful expressions. (Interview 4, 3/8/04)

Furthermore, for Chan, vocabulary, grammar, and usage (collocations or idioms) are all mixed concepts. He said, “if we define grammar as ‘teaching how the language is actually used,’ word usage and idioms are all grammar.” He explained the relationship among vocabulary, grammar and collocation, while drawing an insightful diagram (see Figure 5.5):

We say “idioms,” “idiomatic expressions,” or “grammar.” But they are not separate, but all one. In Korea, we taught “vocabulary” “grammar” and “idioms” separately. However, actually, we have to learn words focusing on expressions, and, in the middle, here is usage or collocation. Let’s take an example of “make use of.” In the past, we taught words and grammar separately. But we can teach them both; if we teach the word “make,” teach like this (circling vocabulary and usage together in the diagram), and if we teach grammar, teach like this (circling usage and grammar together in the diagram). I think this is the main approach of collocation, and if we expand the approach, we can cover all of them: grammar, usage, vocabulary, and idioms (Interview 4, 3/8/04)

![Figure 5.5: Chan’s perception of the relationship among grammar, vocabulary and collocation](image)

On the continuum, Chan was originally in the lexis-focused area. As he appreciated grammar more, then, his position moved to the middle, which is the area of lexico-grammar.
June initially held a clear distinction between words and grammar, while emphasizing the importance of grammar. Later, she still viewed grammar as the overarching concept that included collocation and usage. But she also expressed a more unified concept of grammar and vocabulary:

If we don’t know words, we can’t communicate at all. If we know words, then grammar comes in, and the words are arranged in order and become interrelated. Collocation seems to be the combination of words and grammar. Words have grammar that is influenced by social customs or regulations. I don’t believe that grammarians first generate rules and we follow them. Rather, language has been evolved and changed during a long period of time, and grammarians describe the patterns to show us that some words should be used with certain words in order to make others understood. (Interview 5, 3/10/04)

June held a strong preference for descriptive grammar, and she recognized collocation as the combined area of vocabulary and grammar (i.e., lexico-grammar), which was not evident in her earlier perspectives. Like Chan, she went on to present a diagram that appears in Figure 5.6:

I believe vocabulary and grammar should be combined. I am not sure whether words are included in the big circle of grammar (presented as A in the diagram), or they share a part (presented as B in the diagram), but in either case, usage or collocation is located in the common part. (Interview 5, 3/10/04)
Both Chan and June saw usage or collocation as the area where vocabulary and grammar meet. Here they recognized the role of corpus as a tool for helping them to realize this relationship.

Nick, who had previously been exposed to the corpus approach, stated that “usage is [the] glue between grammar and vocabulary” from early on. He said that his concept of lexico-grammar had enlarged since he started using the corpus. He reaffirmed the idea that grammar and vocabulary are combined while seeing that the same words can have different meanings and can be associated with different grammar patterns:

Actually, rules and words are combined because if there is a word, there is a pattern behind it. Different meanings of the same words have different grammar patterns. This is, you have the same word, two meanings but different English patterns after it. (Interview 5, 3/10/04)
On the other hand, a few participants, Wol and Sung, reported that their understanding of the language did not change much because they used the corpus only for confirming previous linguistic knowledge. Wol said she only checked it for reinforcement, not for new knowledge. Sung responded that before he took the course, he was already aware of the importance of common expressions and patterns in writing, so a new awakening did not take place. It is possible that they did not relate corpus use to their understanding of language because they perceived the corpus mainly as a reference, not as a language learning tool.

**Enhanced collocational knowledge and awareness**

Most participants equated the corpus approach with learning collocation, as the corpus was used to show them the common usage patterns of words (mostly, collocation). As such, they agreed that corpus use increased their awareness of the importance of collocation as well as collocational knowledge of words. However, they did not attribute the enhanced collocational knowledge only to the corpus, but to other reasons as well, including: 1) the writing course, 2) more exposure to English texts and more writing, and 3) their increased content knowledge. While all these reasons were interrelated, it seems reasonable and highly plausible to say that corpus use contributed to heightening their awareness of collocation.

Wol reported that her awareness of collocation increased because “[she] heard and looked at many collocations.” While she was not sure whether this was only because of corpus use or also because of doing more English writing, she commented, “I always
checked the corpus when writing because I felt confident only when I checked the corpus, so maybe I can say [it’s] mostly from corpus use.”

Chan started to become aware of the importance of collocation in the writing course, mostly because of the corpus approach and also through more writing experiences. He said, “in the past, making a sentence itself was a big task, so I didn’t have time to think about more appropriate expressions or usages.”

June reported that multiple examples in the corpus changed her understanding of language because they showed her accurate nuances and usages and helped her with word knowledge. She said “the sense of collocation in writing has been increased mainly because of the class, especially a lot of writing experiences and corpus use. And it could be also because I live here [in America].” She became aware of the concept of collocation when studying for the analogy part of the GRE, but she did not pay attention to the concept after the exam because she did not use it in speaking. However, she began to pay more attention to this as she lived in America and used it in actual communication and writing. But she called for the need to receive corrections on usage errors in order not to repeat the same mistakes.

Nick also recognized his enhanced collocational knowledge, but he gave prime credit to the writing course, and then to corpus use. He said:

Actually, [collocation] was a problem, but I didn’t find an answer. It was a problem for using some words. But now, I find this usage that most people use in the sentence. It was because first I took the classes and then used the corpus. … [Corpus use] helps to make some correlations [between words], make connections and keep in mind connections. I [was] not so aware [of] the importance [of collocations]. And I didn't know whether this collocation can help you to find to be close to native speakers. I wasn't aware of this. (Interview 4, 12/2/03)
In contrast, Sung reported little change in his collocational knowledge:

My sense of collocation has not been increased because I use the similar expressions all the time. I had little chance to find new collocations. I thought I didn’t make any progress in that part. (Interview 5, 3/9/04)

Xiaodong also reported little improvement in his knowledge regarding collocation because his corpus use was so limited. Nonetheless, he said his collocational knowledge was increased to a certain degree due to the corpus-based class activity, albeit through second hand experience:

I have not used corpus so much, but the teacher [has] used it, and other students have used it. They shared the results with us. So, I got something from their results. (Interview 4, 12/11/03)

In general, the participants showed enhanced awareness of the importance of collocation and collocational knowledge. Here it is important to note that when they looked for collocation patterns, they not only focused on forms, but also on meanings. They recognized that patterns using the same words could differ according to nuances in surrounding texts. Collocation is a combined art that intermingles form with meaning. That is, the corpus approach is not only a mechanical process of looking for a form, but it is also concerned with meaning and subtle nuances (“semantic prosody” in technical terms).

Enhanced language sensitivity and discovery learning

Another important influence of corpus use on language learning was increased language sensitivity or language awareness, which was closely related to the above point
about semantic prosody. Once participants were aware of the importance of collocation or common usage patterns, they were sensitized to those patterns. Most of them reported that they began to pay more attention to the combinations of words in their reading as well as their writing. In other words, corpus use not only helped them to solve immediate writing/language problems, but also expanded their language awareness and helped them with their language learning.

Chan reported that he started paying much more attention to the common expressions or collocations since beginning corpus use. His language awareness increased during reading. The following is extracted from his reflection notes:

A change after corpus use (maybe also after the writing course) is that I came to pay attention to the expressions that I ignored in the past. I usually focused on understanding content and didn’t focus on word usage, except for some cases when I found beautiful expressions. However, after the quarter, I came to pay attention to the aspect. The aspect is in two major areas: 1) expressions for which I got corrected or I searched in the corpus, and 2) usage of adjectives and adverbs. In particular, I focus on which nouns/verbs the adjectives and adverbs modify in my reading. (Written reflections, 3/1/04)

In addition to reading, Chan emphasized the synergy of the reading-writing learning process in the awareness and acquisition of the collocation. In other words, his skill sets reinforced each other to firmly establish usage patterns in his cognitive schemata. This confirms Jabbour’s (2001) argument that “[a corpus based approach] provides students with an opportunity to use in writing the word combinations they have come across in their reading” (p.302). Chan said:

Now even when reading articles, I naturally sense the usage. But, I think there is a big difference in getting used to expressions between by reading and by writing. Some expressions, such as “on the contrary,” “in contrast,” are really confusing
even after many readings, but once I use them in my writing, I can remember them more easily. In other words, one-time writing is much more memorable than 10-time reading. (Interview 3, 1/2/04)

In particular, Chan made an insightful point regarding the relationship between awareness and learning. Recall his interview reported in an earlier section where he said that corpus use did not promote new learning, but reinforced prior knowledge. Since that conversation, self-awareness had brewed in his consciousness. He kept asking himself, “what is the actual role of corpus use in L2 writing and learning?” and initiated a discussion with me about the question. Then, in conjunction with his corpus use experience, he found himself paying closer attention, in reading and during writing, to the collocation patterns of words with which he had problems. And this aspect of language awareness suggested answers to himself:

I think my collocational knowledge has been increased. But I don’t think it’s only or directly due to corpus use. Rather, after using it, I started to realize that words have specific collocation patterns and began to notice those patterns. That is, increased collocational knowledge was not the first effect of corpus use, but a secondary effect. (Interview 4, 3/8/04)

In other words, developing the awareness that collocations exist or that they are important in language learning/writing is an educational process in itself. It leads learners to focus on those patterns, resulting in learning. They became independent learners as they transfer the skill to new material and settings. This raises a question about when and how the awareness (exposure and noticing) structures the actual intake and learning. It may be safe to say that language awareness has the possibility to promote a secondary effect and has its own importance and role in language learning.
June paid close attention to usage patterns in her academic pursuits, but also expanded it to her pleasure reading in English. She said:

I didn’t look at expressions in reading in Korean. But, when reading an English text, I focus on each word usage. During the winter break, I read the novel “Big Fish.” I used to read just for getting information, but at that time, I also looked at the usage. I am trying to get more familiar with the patterns. However, although I see those expressions, using them in my writing is a different story. I don’t remember them very well when writing. So, I think we need to use them often in writing to really acquire them as mine. (Interview 5, 3/10/04)

Unlike June, Wol did not consciously pay attention to common usage patterns, although she felt that her collocational knowledge increased as she became more exposed to various English texts.

Wol added that expressions that became familiar through reading naturally came out in her writing. This reiterates a crucial point: exposure and noticing (whether it be conscious or unconscious) can be transferred to learning and writing.

Like Chan and June, Xiaodong also stressed the importance of use in learning collocation. He said that once he used some expressions in writing, he began to notice them in his later reading and found the appropriate usages. More often than not, he learned from his mistakes. In his reading, he became aware that he had incorrectly used expressions in his writing. Those he corrected later in writing. Then, again, corpus use
appears to be self-reinforcing and promotes self-discovery in learning, and learners become independent users of the skill. As Chan explained:

After I write something once, I become aware of it when reading. If I made a mistake last time, I will correct the wrong combination, and next time I will be aware of it more in my reading. (Interview 4, 3/8/04)

In contrast, Sung and Nick ignored collocation patterns in their reading. Sung said he had too much academic reading to do, so he had to skim and avoid paying attention to collocation. It was different, though, when he did non-academic reading, such as newspapers. Nick agreed. He did not pay attention to collocation in academic texts because his purpose was just understanding the content. But in writing, he became more interested in collocational aspects.

The participants’ enhanced language awareness also promoted their discovery learning. In contrast to traditional direct teaching, they used multiple examples or language data (i.e., corpus) to construct common usage patterns and to learn rules about words. The “usage/use” prototype construction discussed in section 5.1.2 showed that students’ corpus analyses were sophisticated enough that they generalized the differences of the two words for themselves. Wol noted the independent learning benefits of corpus in stark contrast to her past English learning characterized by rote learning, i.e., memorizing vocabulary and rules with little reflection. In contrast, she recognized the corpus approach as a tool for more discovery learning. As she examined multiple examples of corpus and constructed prototypes of words, she found some rules herself. According to her:
The most difficult thing in my previous English learning was memorizing idiomatic expressions. We just had to memorize and memorize everything separately with no much connection or meaning. If we had learned English by examples like in the corpus approach, it would have been much easier and long lasting in effect. (Interview 5, 3/11/04)

In sum, participants reported close relationships between exposure, noticing, and learning. Many of them recognized connections to what they had been exposed to previously. What they noticed was later transferred to their learning and writing. The procedural sequence of learning is, though, yet to be determined. That is, it is not still clear whether the students became aware of the patterns first, and started noticing examples during the exposure, or whether exposure to multiple examples led them to become aware of the patterns. However, it appears that the corpus approach heightened their language sensitivity, even if it was a secondary effect, as Chan pointed out, which perhaps can be one of the most important aspects of corpus use in language learning.

### 5.2.2 Changes in the students’ approaches to L2 writing and the writing process

**Perceptions of writing instruction**

Before reviewing the influence of the corpus on writing, the students’ overall perceptions of writing instruction, non-corpus aspects of the class, are presented. This is important because students’ corpus experiences cannot be separated from class learning that helped shape their approaches to writing and the writing process.

Participants unanimously agreed that the writing course was helpful for learning about the overall writing process and the structure of academic papers. The instructor required three major writing assignments that were carefully sequenced. The first assignment was to write a short annotated bibliography of five articles chosen by the
students. The second assignment was a critical review of two or three of the five articles, during which the students could use the annotated bibliography. The last assignment was an expanded paper in which they were encouraged to utilize their previous assignments. As to the structure of academic writing, they learned the basic elements of academic papers (i.e., background, purpose, methods, results, and conclusion), which was often called “BPMRC.” As most participants had few experiences in L2 academic writing, they appreciated learning about the general conventions of English academic writing.

Wol stated that the class helped her to structure academic papers. She employed the structure for writing assignments in other courses during the second quarter. She explicitly differentiated the class’ role and her role as a writer/learner: while the class taught her overall concepts of writing, she was responsible for the acquisition of word expressions or other language-related knowledge:

The class was helpful. Last quarter I didn’t even know how to write, but the class taught me about the structure and organization of academic writing. I am not sure about how much it helped with details such as phrases or expressions, but I think it’s up to the individual student. (Interview 5, 3/11/04)

Chan said the sequenced nature of the writing assignments was what helped him most in the class. This arrangement helped him to organize his writing assignments to become more efficient. He hypothesized that writing a long paper without this approach would have been an overwhelming task. After adopting the procedure (i.e., starting from annotated bibliography, moving to critical review, and then composing the final long paper), he approached writing for other courses the same way. He also said the class helped him by teaching actual textual examples of expressions, e.g., commonly used
phrases in introductions. Thus, it provided a template for future writing experiences. He still felt the need to learn more expressions, but, like Wol, he took personal responsibility for this and did not criticize the course for failing to provide additional expressions. Only he could do the work, he opined.

In this regard then, both Wol and Chan were autonomous learners. They knew what to expect from the class and what they needed to do for themselves. This awareness may have nourished their independence, which prompted more frequent corpus use compared to the others. Another possible explanation might have been that the students’ personal backgrounds tended to influence their perceptions of class activities. That is, both of them had been teachers themselves, so they might have better understood teachers’ expectations and classroom restrictions.

In contrast to Wol and Chan, Sung complained that the class did not explain and teach more expressions. He asked for “more direct teaching about useful expressions as a daily dose.” However, he agreed that the class was helpful for learning how to structure academic papers, although his field – history – does not necessarily follow that structure. Nick and Xiaodong agreed that the class was most helpful in teaching the rules and conventions of writing papers, particularly BPMRC, which provided a core structure in organizing their material.

In addition to learning about the topical structure, June appreciated the class for teaching word choices appropriate for academic writing. She enjoyed learning about the corpus although she did not use it much. She also said the class taught her how to cite sources in academic writing. She commented, “I got a bad grade for my final paper
because of citation problems. I was not familiar with using quotations, and the class taught me how important it is in academic writing.”

Did the students feel their effort in the class was properly rewarded with a commensurate grade? They all agreed in the affirmative. They all believed they received the grade they deserved for the effort they expended in the course. Table 5.4 shows the participants’ writing course final grades.17

<table>
<thead>
<tr>
<th>Name</th>
<th>Wol</th>
<th>Chan</th>
<th>June</th>
<th>Sung</th>
<th>Nick</th>
<th>Xiaodong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td>A-</td>
<td>A-</td>
<td>B</td>
<td>B+</td>
<td>A</td>
<td>A-</td>
</tr>
</tbody>
</table>

Table 5.4: Participants’ writing course grades

Wol and Chan were satisfied with their “A-” grades because although they worked hard, they did not put a priority on the course. Xiaodong, who also got an “A-,” showed less interest in the course:

I didn’t pay so much attention to the class. You know, actually, last quarter I also did [science] experiments. We rotate labs and we have class at the same time. Usually, writing course is not required in our major, so we all think it’s not so important than other courses. (Interview 5, 3/18/04)

June received a “B,” which she thought was reasonable because she had a hard time expressing her ideas even for the last assignment at the end of the quarter.

I didn’t have clear content knowledge, so I couldn’t organize my ideas and didn’t know how to express them in English. Therefore, I had to borrow some texts from

17 Participants’ consent was received for the revelation of their course grades.
references, and in doing so I didn’t know about quotation conventions. I think that was why I got the low grade. (Interview 5, 3/10/04)

Sung received a “B+,” and he believed it was due to his frequent grammatical errors. It was interesting that, as readers will recall, he was concerned about grammar from the beginning and throughout the duration of the study.

The participants’ perceptions of the writing instruction were revealing in that they show what they expected from EAP writing course and how they perceived them relative to their other courses. As beginners in disciplinary writing, they particularly appreciated the lessons about the conventions of academic writing, which prepared them for writing in other courses.

**Perceptions of native-like writing**

The participants were asked about their perceptions of “native-like” writing, as, arguably, corpus is believed to present the most common usages of the language of native writers. Of course, one must ask whose native English do corpora reflect? This is a critical question that often goes unanswered in the corpus literature. In this study, the students referred to a particular variant of English when they referred to “native’s writing.” Possibly it was Midwest American English because that is the English they were exposed to at the time of the study, but it should also be noted that, given their backgrounds, the students had been exposed to other Englishes as well.

One prominent finding was that participants believed that the most conspicuous differences between non-native and native-like writing are the natural word usages and expressions, rather than organization and structure.
Wol believed that non-native speakers can become good writers, but their writing is still different from natives’ writing, particularly in the uses (or lack) of idiomatic expressions. According to her:

Native-like writing is one that employs accurate word usages. Non-natives can also become good writers if they use correct grammar and well organize their thoughts. But we still can tell non-native writing from native writing in terms of usage and idiomatic expressions. They are different in those expressions. (Interview 3, 11/14/03)

In this regard, Wol added, “collocation can be a key to differentiate natives from non-natives.” That is, a good command of collocation is characteristic of native-like writing, and not of non-native writing. Sung agreed that collocational knowledge is the main characteristic of native writing. In his words, “to put the words (that only natives know) in the right place and the right context” is a maker of native-speaker writing.

June pointed out that native writing has “a variety of expressions and idioms to express the same meaning,” while non-native writing often uses “limited sentence structures and expressions.”

Nick was less clear about what constitutes native-like writing. In fact, he was less concerned about what he thought was native-like writing than about “good-enough” writing that communicates content to others.

Xiaodong characterized native-like writing as using common usage patterns (in his word, “habit”) and being simple and clear. Here we can also see his clear distinction between native and non-native writing:

When I read an article, the ideas [are] expressed effectively and [understood] very clearly. This is what I can't find out when I write my paper. When I read an
article, I can understand this is native speaker writing. There are many habit. You
know, “habit.” They express the way we can't find out. That's easy for native
writers to write. For example, I think the sentence is very simple and clear. That
kind of writing is native-like writing. But, for me, I can't write it out when I write
something, because I'm not familiar, because I am not a native writer. (Interview
3, 11/13/03)

Xiaodong, in reference to idiomatic phraseology, added that learning the
combinations of words is very important in writing. He believed that imitation was the
best way to learn it: “We can read many native writing and try to write as the same way.
Use the similar way of the native writers” in order to approximate native writing.

Chan concurred that native-like writing is distinctive primarily in expressions. On
the other hand, he disagreed with the other participants, all of whom did not believe that
non-natives can produce native-like writing. He acknowledged the possibility that they
can. He believed that non-natives can write like natives after extensive practice.

The participants’ beliefs about native-like writing reveal important aspects of their
own L2 writing. They pointed out characteristics that conformed to their own
weaknesses in adopting a “native” style, whatever that actually means. As non-linguists,
they could not detect technical differences with the accuracy of a formal analysis. It
makes sense to conclude, then, that their perceptions reflected the shortcomings they
hoped to overcome in their own language learning. They believed that with practice they
could become good at organization and structure, but they felt they needed much more
training to advance their writing skills in the language domain (i.e., word usage and
collocation). They recognized the usefulness of corpora in acquiring those language
usages. They were aware of the importance of combinations of words or collocation in
writing and saw the value of the corpus approach in overcoming their lack of knowledge in that area.

**Later approaches to L2 writing**

The participants unanimously maintained the idea that writing was crucial for their success in graduate school. By the second quarter, all but June reported that they were more confident about writing than in the previous quarter. They mostly attributed their success to three factors: 1) the writing course (partly corpus use), 2) more writing experience, and 3) their enhanced disciplinary/content knowledge.

Chan found L2 writing much less burdensome than in the previous quarter. He attributed this change to successfully finishing the final long paper in the writing course. In addition, his “bank of expressions” expanded through corpus use and writing practice. Nick also “felt confident in writing after those [three-sequenced writing] classes” although he had no writing project to complete in that quarter. Xiaodong agreed that writing in the second quarter was much easier due to the extensive practice offered in the class.

Wol said her writing became easier in the second quarter mainly due to her enhanced content knowledge. She had switched to a new major, so it was difficult working with the content in the first quarter. The later experience and knowledge gained improved her disciplinary writing in the subsequent quarter. However, while writing may have become easier in general, she was frustrated at making the same mistakes in grammar and expressions. In other words, her intellectual appreciation of writing and approach to writing improved, but troublesome language issues lingered on.
Sung as well felt much more comfortable with writing later in the study. Starting in the first quarter, he constructed a personal database consisting of good sentences taken from his various reading materials. While the other participants also kept note of expressions, Sung was the most motivated to build a systematic text bank. The database was a valuable resource, which he consulted for his writing assignments and used to learn about working with sources. His content knowledge was also enhanced during the study. He became familiar with the jargon and terminology in his field. He summarized his situation this way:

I felt much more comfortable in writing. It’s not because my writing ability increased tremendously all of sudden, but because I learned how to paraphrase texts. I started copying texts for my own purpose and refer to them. By doing so, my skill in using sources increased, composing gets faster, and writing becomes more accurate by using accurate grammar and terminology. (Interview 5, 3/9/04)

June was the only participant who lost confidence in writing later in the study. The main reason was that her major field offered few chances to write in English, thus preventing her from gaining additional practice in writing. In the prior quarter, a seminar she took and the writing course required her to write many papers. The subsequent quarter’s courses were all major-related and did not require much textual writing. She only wrote a few lab reports, and these mostly used mathematical formulae. She felt her writing skills may even have declined because she did not have a lot of writing practice. In her words, “the less practice, the less confidence”:

I didn’t have much chance to write this quarter. If I invested “100” to writing last quarter, I did only about “10” this quarter. So, now I feel my writing skills are even getting worse. I planned to spare some time myself for writing practice even after the first quarter, but it didn’t work out. Without any pressure from outside,
it’s difficult to keep writing. Since I haven’t written anything serious this quarter, I feel very uncomfortable and worried about writing. I have no confidence at all. (Interview 5, 3/10/04)

Overall, using the corpus was beneficial to the participants, as reflected in the fact that they were at a loss when the technology malfunctioned at a later stage of the study. Wol and Sung valued the corpus over the dictionary when they wanted to check about correct usages. June did finally realize that the corpus provided an advantage over other references when it broke down.

Chan was so dependent on corpus that he was helpless when it broke down. Without the corpus, he wrote without consulting a dictionary when he encountered trouble spots. Still, despite his technological dependency, he did not believe that corpus use directly reduced his writing errors:

I don’t think that corpus use directly reduces writing errors. Rather, I think the frequency of errors are related to how much time I spend on revising during the writing process. I think we make mistakes, not because we don’t know, but because we don’t check. When writing, I have a sense that this may be wrong, but I don’t check because I have no time, and then the teacher points out the error. So, corpus helps, but it’s more an issue of whether we have time and willingness to check. (Interview 3, 1/2/04)

Some students noted that reading exposed them to common linguistic patterns that helped them in their writing. Also, Wol developed her reading strategy through practice, and this, in turn, influenced her writing strategy:

I used to focus on every detail of each sentence. But I don’t do it now. I skip unknown words and also skim the texts. I just read the theme sentences at the beginning of each paragraph and skip following examples. Otherwise, I can’t finish all the readings. … From my reading, I learned how to write. Accordingly, I developed a writing strategy. That is, I have outlines on paper by paragraph 1, 2,
and put the theme sentences, and then start filling the paragraph. (Interview 5, 3/11/04)

Wol imagined her professor as her reader, so she tried to write to suit him/her. She said, “when I write, I think my writing is reading for my professor, so I try to write as I read. That is, be clear in introduction, put theme sentences at the beginning of each paragraph, and present a clear summary in conclusion.”

June and Chan also read to improve writing, particularly to acquire linguistic features. June said it was a two-stage process. First, she learned expressions or usages through reading. Secondly, she needed to write them to acquire them as hers’. It was as if the practice fixed it in her mind. Chan agreed that just reading alone doesn’t help much in writing. Retention improved if he used in his writing expressions newly found in his reading. This points to the need for and the advantage of a “learn by doing” or praxis approach.

All of the participants reported that they thought in their native languages while writing in English. However, there was a slow but perceptible shift toward more rapid translation to English. They used their L1s less when employing previously learned expressions, or when writing about simple ideas, while they reverted to their L1s when they had to come up with new expressions or discuss complex ideas. For example, Chan mentioned “I didn’t think in Korean for this writing, but it was because my ideas were not so complicated. If it required more complicated thinking, then I might have used Korean.” In other words, their L1 played two roles: First, it organized the concept; secondly, it was the benchmark for working with unfamiliar expressions.
In summary, as newcomers, the participants struggled with two things: 1) the language issue – adaptation to the English-speaking environment, and 2) academic discourse familiarization – learning a different academic culture and its disciplinary writing practices. In particular, most of them had few experiences in academic writing in English; their English writing experiences were mostly limited to the TWE. As Tribble (1999) argued, writing in a new genre is “difficult” for writers. If we consider disciplinary academic writing as a genre, then learning it was a challenging task for these students at the beginning of the study. Later, their enhanced knowledge about the content in their discipline and participation in the disciplinary environment made their writing easier. However, language issues lingered on, as will be discussed in what follows.

**Later difficulties in L2 writing**

At the beginning of the study, the participants perceived four areas as the main causes of writing difficulties: 1) insufficient command of the target language, 2) organizational and rhetorical concerns, 3) inadequate content knowledge, and 4) unfamiliar writing culture. By the end of the study, they still reported difficulties writing in the first language domain. However, they noted improvement in the other areas and attributed it to writing instruction and writing practice during the first quarter. Therefore, they recognized progress in three of four main problem areas. The language difficulty could not be so easily resolved.

Initially, Wol wanted to be creative. Her goal was to maintain her personal style, her voice, in her English writing. However, her goals became less ambitious as the course proceeded. She decided to conform to the conventional format of academic
writing. In other words, she sought an easy and clear writing style to convey her meaning, an approach that made her feel much more comfortable about her writing. While she still encountered grammatical difficulties, she tried to reduce the frequency of chronic grammatical mistakes. At the same time, she pointed out her lack of a variety of expression and how she tended to “stick to only familiar vocabulary and structures”:

I told you before that I wanted to color my writing and have my own writing style, but once I gave it up, I felt much more comfortable, and the process of writing also got so easy. But I still have difficulties using articles and prepositions. Maybe they can never be fixed completely. I just try to get used to those uses and focus on them in revision to lessen the number of mistakes. (Interview 5, 3/11/04)

Interestingly enough, Wol found that disciplinary writing was easier than expressive writing, such as TWE writing. She asserted that extrapolating from external sources was easier than expressive or creative writing that requires her to initiate her own ideas:

I conducted a case study and wrote about it this quarter. It was easier because I can work with the data, not create something. Interpreting data seems much easier because it is not something that I need to present my own ideas. Last quarter, I had to write my own reflections after reading. It took me much longer time because I need to come up with my own ideas. (Interview 5, 3/11/04)

Like Wol, Chan also commented that writing with sources (either data or reference texts) was easy, but he was not confident about writing without sources or about himself:

Writing this quarter is different from last quarter. Now we write a weekly memo based on the fixed texts. So, writing is not so difficult because it’s not we have to write from nothing. I write based on the texts and borrow expressions from it. So,
I am not sure how well I can do for creative writing. It may be so difficult to write with no text. (Interview 3, 1/2/04)

This is a point inconsistent with what Hinkel (2002) has reported. This study’s findings are generally in line with her argument for the need to pay attention to L2 rhetorical and linguistic features. However, this study’s findings do not support her argument that academic (argumentative) writing is comparatively difficult. She makes a case against the process-oriented writing research that focused on expressive writing, and she argues that expressive writing is easier than the argumentative writing (based on external sources) that most L2 students face in their courses. In contrast, the participants in this study perceived disciplinary writing as easier than expressive writing, which they had concerns about. If these students reflect the norm in L2 learning, then this particular class achieved its goals in preparing students for academic writing. This was clearly not a creative writing class, so the results are consistent with the course’s aims. Hence, it is baffling why students even considered their shortcomings on the creative end.

Chan said his difficulty in writing depended on “the complexity of topic and his thinking.” If ideas were complex, it was difficult to organize his ideas as well as express those ideas in English. June, who reported less confidence in writing in the second quarter, responded that “everything is still difficult in writing: organizing, logic thinking, and expressing….” Sung sometimes did not know how to express his ideas. However, he was creative enough to develop a process to get him through the trouble spots. At these times, he first turned to his collection of language samples to solve the problem:

When stuck in expressions, I look up my own collection of expressions taken from articles. I started collecting expressions in history-related journals since last
quarter. I arranged them by category. It gives me a hint whenever I am stuck. (Interview 4, 2/5/04)

Sung was perplexed by feedback from his teachers. The term “unclear” was especially unclear to him, as it did not direct him toward correcting the problem sentence. But what is interesting is he had acquired the vocabulary necessary to take an analytical approach toward writing. In other words, he had the mental tools to direct him toward possible solutions, as witnessed in this comment:

What I feel uncomfortable now is when I get my paper back and I’ve often found teacher’s comment “unclear.” That “unclear” is unclear to me. I guess maybe I made some mistakes in subject-object choices or words choices. But I don’t know exactly. (Interview 5, 3/9/04)

Sung wanted more specific comments in order to edit his papers and improve his writing skills. Wol also wanted more feedback from teachers, not only content-wise but also in terms of rhetorical issues.

As seen earlier, Nick and Xiaodong were very confident about writing. Nick showed a great deal of confidence in writing, saying that “[writing] was smooth. I was surprised there is no difficulty for me to write a paper.” Throughout the study, Nick attributed his confidence in writing to the writing courses, as reflected in the anecdote below:

At the beginning of the quarter, I wrote an email to my advisor. I said this is my purpose for the quarter to develop this model and get the result. In writing, I put some background, method, result, you know, like we learned from the class. Then, my advisor said, “I like your email. Does your wife help you to write this email?” Because my wife just came here [from Romania] and my writing improved a lot, he thought my wife helped me to write. So, I said, “No, she doesn’t. But I took composition classes.” (Interview 5, 3/10/04)
Xiaodong also did not report any major difficulty in writing. He commented that writing was not so difficult for him, as long as he understood his sources. The main thing for him was to comprehend the articles in order to distill the results in a paper: “the most difficulty is to read articles. Just so many [articles] to read to get information.”

A majority of the participants did not ask for outside help to proofread their papers. Only Wol visited the university writing center. Mostly the center’s counselors focused on basic grammar and mechanics, such as article usages. She did not want them to change her content. The other participants did not have their writing checked by others, nor did they closely check their writing themselves because 1) they usually did not have time, or 2) they were not so concerned about accuracy. The plausible explanations for the latter could be due to limited linguistic knowledge needed for writing in their science major or because of the awareness that non-native writers were expected to make writing mistakes anyway. In the latter regard, Chan did not often carefully check his writing. He knew his teachers were aware that he is an international student, so he expected they were more interested in content and not in accuracy that much. So his writing focused on content, and accuracy was not a high priority for him. The breakdown in the technology was also a factor. He used to edit his writing by using the corpus, but once access to it was lost, he just finished his writing with little revision. That is, the technology breakdown deprived him of a vital resource for editing.

Overall, the participants’ later difficulties in writing were more related to expressions or linguistic issues than to organizational and rhetorical aspects of writing. As discussed in the participants’ perceptions of the writing course, their emphasis on language factors may have been the result of the course focusing on rhetorical knowledge.
of structure or organizational concerns and helping them acquire the general process of
writing.

**Changes in the writing process**

The participants’ overall writing process (drafting, composing, and editing) did not undergo dramatic changes as a result of corpus use. The general process remained the same, but there were minor changes. For example, students paid closer attention to word usage and collocation during writing and editing. While other factors, such as instruction, influenced their writing process, the use of corpus technology influenced the writing process in the sense of enabling them to include a self-editing step. In other words, they took more responsibility for their own language learning and writing. It also gave them more confidence about their writing. Thus, they were able to approach their writing tasks with more confidence and ease. For instance, Wol said corpus use helped her form a habit of checking and double-checking her writing. Previously, she did little revision. Another student, Xiaodong, also commented on the editing step added to his writing process:

> When I write down a sentence, if I have not corpus, I am not sure whether this sentence is right. Without corpus, I will not check after I write all the papers. But with corpus, I will search on that instant in the corpus. (Interview 4, 12/11/03)

Chan also noted the influence of the corpus on his writing process. He often checked the corpus for accuracy and clarity during writing, and so he reported a change not only in the process, but also in the quality of writing. Corpus use thus gave him more confidence in writing. According to him:
My approach to writing did not change all together from a big picture, but since the part of checking the corpus came in my writing process, the quality and process of my writing has changed to a certain degree. In terms of the overall process, my writing process is the same: I finish the whole writing, print it out and revise it. However, I came to check the corpus quite often during writing. I check the corpus to find the most appropriate expressions and depend on the corpus before I produce the final product. If I check the corpus, I feel more confident about the quality of my writing. (Interview 3, 1/2/04)

June reported that the corpus approach changed her understanding of language, which then changed her approach to writing. That is, after she realized that a word changes meaning depending on the context, she began to pay more attention to the uses of words when writing.

Nick believed the class overall changed his writing process by teaching him how to structure an academic paper. Using the corpus, though, increased his confidence in writing “because I know I have 40 sentences I can pick the best phrase, and I feel very comfortable because I know it’s right.”

To summarize, the corpus approach changed the participants’ understanding of language: it promoted their perceptions of lexico-grammar, collocational knowledge, and language awareness. The change, in turn, was reflected in their approaches to writing and writing process. While some participants used the corpus more than the others, their exposure to the corpus approach made all of them aware of the importance of common usage and collocation in writing, and they came to focus not only on content, but also on the linguistic domain. In other words, the corpus served as an actual reference, and also served as a catalyst to promote the awareness of the importance of collocation.

Corpus use also affected the participants’ approaches to writing and the writing process. Once the corpus approach was introduced to the writing process, they took more
responsibility for their writing by checking the corpus, while paying more attention to linguistic matters. By doing so, they approached writing with more ease, and their confidence in writing increased.

5.3 Students’ perspectives on collocation in L2 learning and writing

Collocation is at the core of lexico-grammar in the corpus approach. As such, the participants recognized the importance of collocation in L2 writing through their writing and corpus experiences. In this last section we present the students’ beliefs on the role of native language in L2 collocation uses and their views on collocation in L2 learning and writing.

5.3.1 Role of L1 in L2 collocation uses

Most participants believed that their L1s interfered with their L2 collocation uses because each language has its own collocation patterns. Wol said Korean was helpful only for concepts that have similar expressions in English. For example, the Korean concept for “tie up one’s waistband” (허리띠를 줄라매다) is the same as “tighten one’s belt” in English. Thus, according to her, the key for positive transfer from L1 to L2 expressions was to know the similarities and differences in usage between the two languages:

Korean is sometimes helpful and sometimes unhelpful. It seems languages have universal usage or maybe it’s just because modern Korean is heavily influenced by English. I’ve often found the two languages have the same expressions. For example, for “tighten one’s belt,” we say exactly the same in Korean, right? I found many cases like this. We have the same ideas and expressions. If they are different, that’s the case when I make mistakes. So, we need to know which are the same and which are not. (Interview 5, 3/11/04)
Wol added that context is the key to knowing whether a collocational pattern fits the meaning. She said: “even for the same way of thinking, actual expressions involve the matter of collocation in English, and for that matter, Korean doesn’t do much. Because collocations are unique in each language, we have to consciously memorize them.” Chan gave an example of incorrect collocation uses in English due to the direct translation from L1:

I used “trial” to express “attempt” in my writing assignment. “Shi-do”(시도) in Korean is “trial” in English, right? But, the teacher pointed it out and commented that trial is used when something didn’t go well, so I have to use “attempt” or “effort” to mean I put some effort in it. “Trial” “attempt” and “effort” are all translated as “No-ryuk”(노력) in Korean. But, in order to decide which is appropriate in the context, we need to get used to usage patterns in English. (Interview 2, 11/10/03)

Chan went on to provide an interesting comment about the relationship between thoughts and languages. He believed that thoughts are not either in L1 or L2, but exist as an “image” that can be expressed by known expressions. For him, familiar concepts are expressed easily also in English, like “apple,” but he had to translate unfamiliar words and expressions from L1. In other words, if his repertoire of English expressions contained what he wanted to express, he could use English directly. Otherwise, he had to utilize his L1 to express his ideas, which most likely caused errors in English writing. He said:

What we think is more of “image,” not language. It’s not in English or in Korean. It’s just an image. For example, we know “apple” by image before we come up with the word, “apple,” either in Korean or in English. When I speak Korean, the image just comes out in Korean right away. That’s because my image and language are related. When the image is simple, I can express them in English as well. When we look at an apple, I just come up with the word “apple” in English.
I don’t think too much. If I have a big bank full of collocation, I can just pull out the expressions I want. I don’t need to rely on Korean. But, when the image is complex, I first need to use Korean to express it. I translate Korean into English because I can’t directly express the image in English. I often have to use Korean first, then use English, which usually causes errors. It seldom helps. (Interview 4, 3/8/04)

June also reported unequivocally that Korean interfered with her English writing. She strongly believed that L1 interfered in three respects: 1) time, 2) difference in overall logical structure and sentence structure, and 3) nuance and expressions. She thought in Korean first and then sought corresponding English words, which meant more time was required to complete what she was writing. Secondly, she complained about the different linguistic characteristics in logic and sentence structure:

The two languages have different grammar. They are different in overall logical structure or even sentence-level structure. So, the Korean way of thinking interferes with English writing. For example, in Korean, verbs come last, and the most important information comes first. Therefore, we often have a long subject. But, in English, we need to have a short and unheavy subject, while putting all the important information later in the sentence. (Interview 5, 3/10/04)

As for nuances and word combinations, June’s experience in Korean helped her to understand that English has its own collocation patterns as does Korean, so she was careful in word choices. In another sense, corpus use created awareness of collocational patterns in the L1 as well, which is an added benefit. However, except for that understanding, the Korean experience was not beneficial for actual usages because “words in the two languages are not in one by one correspondence.” When she referred to Korean for corresponding English words, there were too many possible choices, including different nuances and collocation patterns. Thus, attempts to transfer often
caused errors. June pointed out, in particular, the difference among word classes. She said, “noun or proper noun may be ok, for example, banana, because there is only one correspondent word. But, verbs, adverbs, and adjectives are particularly problems.” In her own parlance, then, she expressed the linguistic principle that lexicalized items that have unambiguous referents can be correctly translated into other languages, but delexicalized words that have obscure (or several) meanings are particularly problematic in translating word usage.

Sung agreed that Korean linguistic patterns interfere with English collocation uses and writing because “if [he] expresses in a Korean way, the meanings change in English.” Most of all, he complained about different word orders between the two languages. Another student, Xiaodong, indicated that Chinese did not help him with English expressions and writing, although it was helpful for outlining writing at the conceptual level. He said:

Chinese expressions and combinations of words [are] not helpful for English. If you always think about Chinese when writing the detail, then it will be Chinese-English. So, for details, it’s not helpful. (Interview 5, 3/18/04)

On the other hand, Xiaodong commented that his learning helped him to discriminate between the two languages’ different usages. Chinese expressions are different from English, but the differences did not always translate into errors in his English writing because he was aware of the differences. This also means that if he had not learned the differences between L1 and L2 usages, he was likely to err in those usages. Nick made a similar statement. Learning collocational uses helped him to avoid errors in English usages. For example, he learned that “discuss about” was incorrect in
English, while it is correct in Romanian. However, he reported that there was not much difference between his L1 and English. In fact, he was the only participant who strongly believed in the positive effect of L1 on English learning and writing. According to him, Romanian is similar to English, so it helped him with English learning.

In sum, most participants reported a negative effect of L1 on L2 collocational uses, as each language carries unique collocation patterns. More important, though, the data indicated that the participants were careful observers of their L1s and L2. They demonstrated a sophisticated understanding of the differences in the two languages as a result of their experiences and learning. As users of two languages, they were aware of contrastive linguistic and rhetoric features without knowing related theories. Based on their own observations, they adapted their use of their L1s in their L2 writing. They knew when to rely on and how to use their L1s as resources in L2 writing. It is also noteworthy that their learning helped them become aware of the differences between languages and helped them to avoid errors.

5.3.2 Collocation and L2 writing

The participants uniformly agreed that collocational knowledge was more crucial for L2 production than for comprehension. They perceived that reading and listening do not require extensive collocational competence in understanding external content. In contrast, writing and speaking were perceived to demand proficiency in use of word combinations, given the nature of such production. In particular, they believed that collocational knowledge was critical in writing to achieve more efficient and competent conveyance of meanings. They regarded correct uses of collocation as the epitome of an
advanced level of English and writing proficiency. In other words, it was the benchmark for total fluency, in their estimation. For example, Wol said, “competent use of collocation indicates great familiarity with English writing.” Chan also said:

I think collocation is very important in writing. When we use the expressions that do not match, it may sound very awkward and uneducated, although meanings can come across. (Interview 4, 3/8/04)

As noted in an earlier section, the participants perceived collocation as the key factor of native-like writing. Wol said, “collocation can be a key to differentiate native from non-native writing.” June associated collocation with cultural knowledge as well as native-like production. For her, “natives just feel whether expressions sound natural or strange. But I can’t, because I didn’t live here and I just learned English from school.” She emphasized the importance of cultural familiarity and natural acquisition in order to increase the knowledge of word combinations.

June and Xiaodong believed that collocation played a particularly pivotal role in formal writing because of its emphasis on conventions and formality. According to Xiaodong:

I think when you’re writing a paper, it’s more important. You need to pay more attention to the combinations. But in your everyday life, for example, just writing an email, I don’t think it’s so important. We just need to understand the email. (Interview 5, 3/18/04)

The participants believed that the best way to improve collocation knowledge was through repeated exposure through reading and practice through writing. Wol said, “we need to read and write a lot. If we read a lot and get used to the patterns, we can use them
naturally.” Chan said, “we need to use collocations despite making mistakes and get them checked.” June agreed that “we should pay attention to the patterns and use them in writing.” On the other hand, she wondered “whether I need to memorize one by one, or if I know a certain rule, I can get other patterns naturally.” This indicates that she perceived collocation from a rule-based point of view.

Sung and Xiaodong specifically pointed out the value of imitating native collocational uses in writing in order to increase their repertoire of collocation. Sung believed that copying collocations from reading and then applying them in writing may be simple, but it is the most efficient approach. Xiaodong said, “we need to read more papers and try to write the same way and use the similar way the native writers have used.” In particular, he stressed the importance of writing practice for acquiring word combinations. According to him, “the more you use them, the more you will learn the combinations. You just write them many times, and then you will be used to this kind of collocation.” This view expresses his belief that using collocation in writing would increase awareness of the patterns, leading to conscious learning.

The participants were also asked about the necessity and characteristics of collocational knowledge in their own disciplines. Sung and June acknowledged the need for collocational information specific to particular disciplines and asked for a specific corpus for their disciplinary areas. For example, June said:

I hope the corpus can have a function that I can choose newspapers, or scientific journals or my major-related category so that I can search only those texts. It may not help me to learn general English, but at least, it can help me with my writing that I need to produce. (Interview 5, 3/10/04)
In contrast, the other participants reported that general English collocation was sufficient because their disciplines did not require distinctive collocational knowledge. They believed that general collocation knowledge can be useful in other areas of writing as well. Nick said:

I think [collocations] are applied generally. It’s not only for one field. They’re applied generally to be able to develop paper, or to develop email, to develop composition. Not only for my field. It’s general for writing, and collocation provides you the usage, and this is the best thing that it has. (Interview 5, 3/10/04)

While the participants agreed on the importance of collocation (and thus its teaching and learning) in L2 writing, they addressed reality constraints in classroom implementation in their home countries, based on their previous learning experiences. Before they came to America for graduate study, all of them learned English in EFL situations. As EFL learners who were rarely exposed to English in everyday life, they were sensitive to the differences between ESL situations and EFL contexts.

Nick believed that the best way to teach was to present general rules first, then details later in EFL contexts. That is, “first rules, then usages.” Interestingly, while he complained about his previous learning experiences at the beginning of the study, he still valued that type of teaching. In other words, he held the view of language learning that was formed during his previous learning experiences, as opposed to his later view that grammar and usages were combined, and so should be presented together. His comments about contextual concerns offered one possible explanation. He differentiated the pedagogical needs between teaching ESL and EFL. He believed that general rules should be taught first in EFL settings, while ESL situations can start with teaching collocation:
It’s better to learn rules first and second step, start more details. When I said general rules, I think somebody [who] starts to learn English not here. If you are here in OSU, students here, I think, you pass that part. Now you need to learn more collocation and more detail. But, when you start from the beginning, you are a beginner. You need to follow the rules to understand the [tenses] and articles. This is my thought about the stage. (Interview 5, 3/10/04)

June believed that classroom learning in the EFL context was artificial, since it lacked cultural contact and natural learning opportunities. Thus, she noted the need for teaching grammar rather than collocation, which requires natural acquisition:

I think class learning is different from acquisition that occurs naturally by living, because in class learning, language is presented as knowledge, not as an actual communicative language. Class teaching can be time efficient, but it’s different from the natural language learning process. Nonetheless, I believe we should teach grammar first in EFL situations because there we can’t encounter and use the language in real life anyway. (Interview 5, 3/10/04)

Wol pointed out the limited language focus in her English learning. She complained that her previous English education emphasized fixed idioms that were infrequently used, while ignoring phrasal verbs that were more frequently used. She said:

If we learned by examples like the corpus approach, it would be much easier to learn. One ESL teacher lent me a book about phrasal verbs. The book has a lot of examples about usage of basic verbs, such as “take.” I found it much more useful than idioms that are not commonly used in communication. (Interview 5, 3/11/04)

Wol’s observation was an insightful point regarding collocation restrictions that the literature has also addressed. For example, Howarth (1996) presented collocation in a continuum that has free combinations (e.g., eat banana) on one end and fixed idioms (e.g., foot the bill) on the other, according to the degree of restrictions in word combinations. He argues that English teaching has neglected the less restricted collocations, such as
phrasal verbs and delexicalized verbs, while it has focused mainly on highly restricted idioms.

There was another side of the argument which was worth noting. Sung did not present a clear idea about teaching collocation, but he commented, “maybe we don’t have to teach collocation separately, as we didn’t.” This can be interpreted as his satisfaction with the traditional way of learning.

In contrast to the other participants who were cautious about integrating the corpus approach and teaching collocation in EFL situations, Xiaodong strongly believed in the need for and potential benefit of collocation for his home country. He believed that Chinese students needed the corpus more than English learners in America because of the lack of resources for learning English in China: “I think students in my country need such a thing more. They should have tools to improve their English.” This is consistent with Conrad (2000), who argued that corpora can improve language teaching in foreign countries, despite the lack of native speakers in the teaching context.

In sum, the participants recognized the importance of collocation in writing in general and formal writing in particular. They also indicated awareness of a mutual relationship between collocational knowledge and writing. That is, collocational knowledge can improve writing, but at the same time, writing can enhance collocational familiarity by increasing awareness. On the other hand, they addressed reality constraints in teaching collocation in EFL contexts. They believed that since only minimal language use was possible in such contexts, rules should be given more attention. While they had complaints about their previous learning experiences at the beginning of the study, it is worth noting that they did not entirely devalue those experiences later. On the contrary,
they appreciated the way they were taught. That is, the students were able to perceive contextual differences and presented different needs and teaching methods accordingly.
CHAPTER 6

DISCUSSION AND CONCLUSION

This chapter summarizes the study’s major findings and provides implications for pedagogy and research. Section 6.1 is a summary of students’ corpus use experiences as related to the research questions that launched the study. It reviews the benefits and influence of corpus use on L2 writing, illuminating the role of corpus technology in L2 writing. Section 6.2 follows with suggestions on how to integrate the insights from the study into L2 writing instruction. Lastly, Section 6.3 presents the limitations of the study, followed by important recommendations for future research in Section 6.4.

6.1 The role of corpus use in L2 academic writing

The primary interest of this dissertation was to answer the overarching question, “In what ways does corpus technology facilitate L2 academic writing (instruction)?” The research questions posed in this study were:

1. How do the students use corpus technology in L2 writing?
2. How does corpus use mediate the students’ understanding of language and L2 writing?
3. How does the native language influence the students’ L2 collocation uses? What role does collocation play in L2 learning and writing?

The research site of the study was a graduate level ESL academic writing course at a large American research university. While data were collected from all students in the class, six students participated in an in-depth case study and became the main focus of the research. This study primarily adopted a qualitative framework in order to obtain rich descriptions of corpus pedagogy in L2 writing. The data came from a wide array of sources: 1) classroom observational notes, 2) interviews, 3) recall protocols, 4) think-aloud protocols, 5) corpus search logs, 6) class corpus search assignments, 7) written reflections on corpus use, 8) SAS inventory, 9) SILL inventory, and 10) questionnaires.

This study constructed portraits of the focal students consisting of individual life experiences and writing experiences. Those experiences were a lens through which to understand different contexts of the individual students’ L2 writing experiences. The qualitative/ interpretive approach helped to identify the students’ individual needs and explore the difficulties they faced in adopting the corpus technology.

6.1.1 Students’ corpus use in L2 writing

The first research question explored how the students used and perceived the corpus in L2 academic writing and what were the individual experiences and learning contexts affecting those uses and perceptions.

The study approached this question by examining the six focal students’ corpus use patterns and self-evaluations of this use. The students used the corpus as a linguistic resource to quickly check their writing; they checked the search results and thus solved
their own immediate problems. They preferred to use the concordance (sentences) data over the collocate (frequency) data because the latter tended to take much longer to produce, and the students wanted to look at how words were actually used in sentences. On the whole, they reported that corpus use helped them to see the contexts of word use. That is, it showed them “how related words are used in different ways and are appropriate in different contexts” (Biber et al., 1998, p.21). The main linguistic elements they checked were: 1) prepositional usage, 2) verb + noun connections, 3) delicate meaning differences or nuances, 4) correct uses of words and contexts of use, and 5) proper modifiers. This suggests that students developed a new awareness of lexical use: that words do not have exact meanings, but require contextual interpretations (McCarthy, 2001).

Some participants used the corpus for initial composing of their writing assignments, while the others used it mainly for revisions. Interestingly, the more frequent users, Wol, Chan, and Sung, used it for initial composing, as they consulted the corpus while writing whenever possible. On the other hand, all of the participants reported that corpus use helped them with sentence-level composition and revision rather than in global content or at discourse or organizational levels. This perception may have been related to the characteristics of the corpus used in the course, that is, the general corpus that was not created for specific academic disciplines. A more likely explanation is the influence of the instructor’s pedagogical model for the design of the course. Students were expected to use corpora to solve their sentence-level writing problems for themselves, while the teacher used other materials (e.g., process-oriented class
assignments) and activities (e.g., constructing a style manual for academic papers in a given field) to teach organizational and rhetorical aspects of writing. At the end of the course, the combination of corpus use and classroom instruction seemingly contributed to the students’ successful writing experiences. The students perceived the corpus mainly as a valuable resource for addressing their linguistic problems. This is in line with Tribble’s (1999, 2002) point that corpus linguistic tools and resources are particularly useful for acquiring linguistic and contextual knowledge. According to Tribble, effective writing production requires four types of knowledge: content knowledge, writing process knowledge, context knowledge, and language system knowledge. He emphasizes the potential benefits of the corpus approach to EAP writing instruction, which emphasizes lexico-grammatical and contextual analysis. This type of analysis can help students extend their understanding of academic literacy.

The think-aloud protocols of prototype construction revealed the participants’ capability as sophisticated corpus analysts. The participants identified unique features of words, categorized them, and finally constructed common usage patterns from corpus data analysis. Their processes reflected Johns’ (1991) description of the basic procedure that was involved in concordance-based learning: “identify-classify-generalize.” One interesting finding from the think-aloud task observations was that they all first and readily recognized the usage patterns they learned in their previous English education; they recalled and confirmed their prior knowledge with little doubt about it. However, a few participants’ analyses did not extend to new knowledge beyond the reconfirmation of prior knowledge. This raises an important question as to the role of corpora in language
learning: whether corpus use corrects, confirms, or reinforces students’ prior knowledge, or whether it extends the process to new learning. The data indicated that it varied, depending mainly on the purpose of the students’ corpus searches, their levels of analytic skill, and the depth of their resultant interpretations. More often than not, the participants used the corpus to reinforce and confirm their prior knowledge rather than to generate new knowledge. This was more evident in searches for explicit grammatical points, such as prepositional usage. Corpus use led to new knowledge, however, when users searched for the most frequently used proper words in certain contexts and tried to discern subtle nuances between similar words. This validates the usefulness of the corpus (lexical) approach that emphasizes the co-occurrence and typical contexts of words, i.e., providing collocation information (Biber & Conrad, 2001; Conrad, 2000; Jabbour, 2001).

Generally speaking, the participants viewed corpora as a useful reference for linguistic help. In comparison with dictionaries, a time-honored traditional language resource, the participants used the dictionary for definitions and corpora for usage. While they perceived the dictionary as a tool for learning immediate meanings of unfamiliar words, they viewed corpus use as a second step in solving their usage problems, i.e., as a linguistic checker. As time went on, the participants used the corpus more frequently than the dictionary. As advanced learners, they needed more usage knowledge than word meanings, and so their frequency of corpus use increased. Interestingly, they did not revert to their dictionaries even when a technology breakdown made the corpus unavailable. Corpus use had taught them that the dictionary could not solve problems related to relationships between words. As time went on, the students creatively
integrated the new technological resource with the conventional reference tools. In other words, the students’ experiences taught them the capabilities and limitations of each resource, and they adjusted their uses accordingly. This challenges the techno-deterministic paradigm which asserts that new technology will replace traditional methods in an evolutionary sequence. The results agree with Tribble (1999), who argued that corpus resources could be used as a complementary method for L2 instruction, not as a replacement for traditional, or more familiar, teaching methods.

Corpus use provided the participants with not only textual help but also with a psychological advantage in writing. The corpus instructed them on correct common usage, and it enhanced their confidence about their writing. This indicates a psychological, emotive aspect to learning, as learning requires a broad cognitive approach taking into account affective issues as well (Gardner & MacIntyre, 1992).

While the corpus provided the students with many benefits, the students also encountered difficulties using it. They commonly agreed on its advantages (i.e., showing common usage patterns and contexts of uses), but their difficulties diverged, ranging from technological aspects to time availability.

The students agreed that availability/portability was a decisive factor in their corpus use. This factor includes not only whether they had a computer or Internet connection, but also whether the corpus was available at the time they sat down to write. Some of them complained about slow Internet speed. Quite a few students pointed out the time-consuming nature of data searches and analysis. Although the corpus is often described as user-friendly and easy to use, students’ use was limited by difficulties in
using searching query functions and analyzing the collocate data. The difficulties were partly due to their high expectations for acquiring advanced-level corpus search beyond the simple search. Here, we must heed Ferris and Hedgcock’s (1998) admonition about use of another technology, word processing, in L2 writing: “ESL specialists caution that students must be given adequate instruction and assistance in using word-processing tools for them to be more help than hindrance” (p.274). Likewise, in the case of corpora, the technology alone is no single answer to improved learning. Training is required for students to learn how to use the technology in order to take full advantage of the resource.

In addition to the technology-related difficulties, a concern was raised regarding using the general corpus that was not created for specific academic fields. One student, Sung, stressed the need for a specialized corpus in his field, history, while the others were mostly satisfied with the general corpus used in the course. There is no question about the distinctive contributions that specialized corpora make in L2 writing. But, it is also useful to recognize the general corpus’ strengths. Here it is important to recall that discourse level or rhetorical characteristics can be acquired more quickly than linguistic features by advanced L2 learners. Specialized corpora are usually believed to be useful for teaching a certain discourse pattern and the language commonly used in a specific genre. This study showed that this aspect can be dealt with by using alternative materials associated with the disciplines. In contrast, general corpora can be used to help students deal with their chronic, harder-to-acquire linguistic problems on their own. In particular, as seen in the participants’ behaviors in this study, students can learn technical words
specific to their disciplines through their own reading and content knowledge. This is consistent with Jabbour’s (1997) point that the specific terminologies which occur relatively infrequently in student readings, can be easily handled through references, with minor help through instruction. She adds that students may be more familiar with the subject words than their EAP teachers, while those words may be irrelevant to the class as a whole. In sum, general corpora that present the most frequently occurring general words can be used effectively, because they cater to the needs of all the students in an EAP classroom.

The across-case comparisons as well as the within-case analysis showed that individual experiences and contextual factors affected the participants’ uses and perceptions of the corpus in L2 writing. The individual and contextual portraits of the students set the stage for interpreting their writing and corpus experiences. Although the participants were in the same writing course, they reported different issues and problems in corpus use and writing. These include technological and analytical skills, disciplinary content knowledge, English proficiency, and writing proficiency. This range illustrates the very complex nature of teaching L2 writing.

The across-case examination initially identified three pairings of the six participants based on their perceptions and experiences in L2 writing and corpus use. For discussion, the participants were paired into three groups: Wol & Chan, June & Sung, and Nick & Xiaodong.

The first pair, Wol and Chan, generally enjoyed writing, were very experienced in L1 academic writing, and were also confident about L2 writing. As they had a genuine
interest in L2 writing, their goal in writing was not just to communicate ideas, but also to compose persuasive, powerful texts, and to develop their own writing styles. With respect to the corpus, they were initially positive toward its use, as indicated by their frequent corpus use. In addition, the data showed that Wol and Chan were autonomous learners and independent corpus users. They often consulted the corpus for their own needs beyond the class requirements. It is clear that their genuine interest in writing development prompted more frequent corpus use and nourished their independence compared to the others.

In particular, Chan’s corpus use deserves discussion. As the most enthusiastic and frequent corpus user, his corpus use illustrated a wide range of corpus searches and considerable depth of analysis and interpretation. The other participants’ corpus searches were mostly limited to so-called “substantial” linguistic elements for sentence construction, such as verbs, nouns and prepositions. Chan had a similar repertoire at the beginning, but expanded it to include “less crucial” linguistic features (e.g., adjectives and adverbs). The fact that his searches gradually narrowed to mainly modifiers indicates his sophistication as an extensive corpus user.

The second pair, June and Sung, disliked writing. Both were inexperienced L1 academic writers, but Sung did have experience in learning L2 academic writing in Japanese during his master’s studies. In general, they lacked confidence about L2 writing in English, and their primary concern was to get their ideas across on paper, with an emphasis on correct use of linguistic features. They initially showed a negative attitude toward corpora, albeit for different reasons: June had difficulty using the search
technique, and Sung questioned the usefulness of the corpus over the dictionary. Their attitudes, however, became positive, mainly because of increased use, which thus exposed them to its usefulness. June appreciated the weekly corpus search assignments that gave her more opportunities to practice. This suggests that students may need to be exposed to technology for a certain period in order to effectively use it or like it.

Sung, the only participant who initially showed a clearly negative attitude toward corpus use, turned out to be one of the most frequent corpus users. This was an unexpected finding. His studies in history required a great deal of writing, so he had a need for and a keen interest in writing improvement. He also expressed the need for a specialized corpus in his field, even though it would appear that history would require less technical, more generalized corpora.

June’s change to a positive attitude did not prompt her to use corpora more frequently. In fact, June struggled with corpus use, which was related to many issues. One reason was that she needed more training to use the search technique effectively. Second, her lower level of language proficiency hindered her corpus use. She had a hard time selecting words to begin her searches. She had to consult a bilingual dictionary to find English correspondents before engaging the corpus for word usage. Also, analyzing the corpus data was not easy for her. She had to examine, categorize, and translate the data into prototype strings. To effectively use the corpus, one has to reach a certain level of language proficiency for efficient searching for and analysis of the data.

Another possible explanation could be June’s past learning experiences. As she was taught previously by a traditional transmission model, her corpus use might have
reflected less independent learning or learner autonomy. She wanted ready-made results from others rather than to analyze the data herself. Recall, however, that Wol and Chan were taught the same way June was – transmission model – but they were autonomous learners and used the corpus frequently. While their genuine interest in writing improvement increased their corpus use, this raises an important question about the relationship between autonomy and inductive strategy. Some corpus researchers (Johns 1991, Thurstun & Candlin 1998) have argued that corpus use enhances learner autonomy through an independent problem-solving approach. What is not clear, though, is whether corpus use enhances independent learning, or whether only already autonomous learners are attracted to corpora.

The last pair, Nick and Xiaodong, did not like writing in general: they wrote only because they had to do it, even though they were confident about L2 writing. However, both of them enjoyed emailing and web-based chatting with others. Likewise, their primary focus in writing was to communicate content to readers. In another sense, they did not care so much about overall quality of text as long as the writing was “good enough.” Initially, they were positive about corpus use. In particular, Nick, who had been exposed to the corpus for almost one year, was highly enthusiastic toward the corpus. Interestingly, though, Nick and Xiaodong did not use the corpus often during the study. Xiaodong attributed this mainly to his lack of time, while Nick had fewer writing assignments with which to use it.

Nick’s corpus use deserves some separate consideration. He did not use the corpus as much as indicated by his professed enthusiasm for it. In addition, the think-
aloud task for prototype construction suggested that he did not have an efficient or refined strategy in his analyses. Certainly, a one-time task observation is insufficient to fully capture his strategy toward or ability in corpus analysis. However, a student’s purpose for corpus use seems to be related to the sophistication of corpus analysis and interpretation. Nick clearly stated that he used the corpus mainly to find prepositional usage, which may not require a sophisticated search and analysis, but can be obtained by a quick look at the data. One primary reason he infrequently used the corpus was due to his science-related major, which required few writing assignments. In fact, he simply did not have to use the corpus. In addition, his increased confidence in English writing required less use of the corpus as time passed.

To reiterate, two pairs (Wol & Chan, Nick & Xiaodong) were similar within the pairs and showed differences between the pairs in their corpus use behaviors and perceptions, while the third pair, June & Sung, showed differences between them, particularly in terms of the frequency of corpus use. On the whole, Wol, Chan, and Sung used the corpus much more frequently than Nick, Xiaodong, and June. The former three used it for other courses and their own needs in L2 writing, as well as for the writing class, while the latter three only used the corpus for class writing assignments. One possible reason for this difference is the field of study. The latter were science majors, where they may have had less need for writing in their courses. In contrast, the former were non-science majors (i.e., education, history, and science education), all likely requiring more writing, and of a wider varieties of styles.
Also noteworthy was the relationship between participants’ frequency of corpus use and the number of writing assignments, closely related to the above point and the characteristics of their field of study. The amount of writing tasks varied among the students, and this affected their corpus use. Most of the participants used the corpus more pragmatically as a problem-solving tool in writing: their corpus use was need-based. Thus, their corpus use decreased when they had fewer writing assignments. The three students, Nick, Xiaodong, and June, who did not use the corpus much throughout the study, had fewer writing tasks to perform and thus less need for corpora. However, they indicated that they planned to use corpora more when they encountered more writing assignments and formal papers in the future. This need-based attitude toward corpus use is a revealing finding. It suggests that technology use is related to the actual needs of the users. If the technology is not geared toward the students’ immediate needs, it may not capture their interest despite its value in improving their writing. Hence, despite the usefulness of the technology, lack of meaningful engagement with writing would limit its appeal and therefore its use.

As seen thus far, a wide array of individual experiences and learning contexts were involved in the students’ corpus use and their successful experiences with the corpus. In this context, Turnbull and Burston (1998) have pointed out the complicated nature of students’ corpus experiences:

Language learners are likely to experience varying levels of success with concordancing strategies and that concordance data alone does not necessarily stimulate inductive learning strategies. A continuum of abilities and familiarity with such strategies exists, and is very much influenced by previous experience of language learning and individual learning styles. (p.10)
While little research has examined the variables mediating the effect of corpus use on language learning, it has been proposed that the effect may be different depending on individual differences, such as the learners’ level of language proficiency and learning styles (Johns, 1991, 1994; Stevens, 1991). This study used the SILL and SAS inventory as well as students’ self-reports to examine the students’ learning styles. The results of the standard inventories did not identify distinct characteristics among the students. As advanced graduate students, they probably shared common learning styles by background or from successful learning experiences. On the other hand, the self-reports explored the unique attitudes of each learner while reflecting a more dynamic learning process than the standardized inventories. Given the study’s small sample size, we cannot conclude that self-reports are generally more useful than inventories for identifying students’ learning styles, but it remains an open question whether the standardized inventories capture all other relevant characteristics, particularly when cross-cultural values are at issue.

In short, this study found that, at least for these six advanced learners, learning styles did not have a great impact on corpus use. Rather, individual experiences and other contextual factors played a more important role in determining the frequency and range of corpus use, selection of linguistic items for searches, the depth of analysis and interpretation, and thus the successful exploitation of corpus resources. The experiences and factors identified in this study were various: characteristics of the field of study, goals of writing, interest in writing improvement, needs for the resources as related to the number of writing assignments, familiarity with the corpus technology, time availability,
level of English language and writing proficiency, and nature of writing experiences in L1 and L2.

6.1.2 The influence of corpus use on L2 learning and writing

The second research question explored the role that corpus use played in the participants’ understanding of the language learning and writing processes. Most of them perceived the corpus as a valuable linguistic resource to help them to become more proficient in L2 writing. However, the study found that when the students accepted the new technology into their repertoire of L2 learning and writing, its effect went beyond its usefulness as an immediate linguistic checker. This study revealed the various roles corpora can play in language learning. One role is to increase awareness of word usage in contexts, rather than over-emphasizing “single words out of context” (McCarthy, 2001, p.63). Another role is to provide multiple, representative examples of language in use. A third is to expand the understanding of functions of linguistic items by focusing on forms as well as meanings (Jabbour, 2001). Still a fourth is to promote discovery learning in the learning process (Johns, 1991).

With respect to understanding of the target language, most participants expressed an integrated concept of vocabulary and grammar, or “lexico-grammar” in technical terms, by the end of the study. While it is not clear whether this was solely the result of corpus use, the corpus approach that focuses on collocation and that intermingles form with meaning pushed the boundaries of grammar and lexis to a merged area and promoted positive perceptions of lexico-grammar. While the participants’ pervious training mainly emphasized grammar, during the study they became aware of the value of
the lexical approach (Lewis, 1997, 2000). As advanced ESL students, their writing problems were often related to word usage rather than grammatical aspects, and their experiences with corpora enabled them to address these problems meaningfully.

The traditional linguistic approach mainly focuses on the structure or rules of the language. In contrast, the corpus linguistics approach looks at the most frequently used words, their common usage patterns, and collocations. Lewis (2000) argues for replacing the traditional emphasis on grammar with an emphasis on lexis in language teaching. However, his proposition does not exclude teaching grammatical aspects of a language: rather, it locates grammar inside the lexis. According to Lewis, “Recognizing that every word has its own grammar means that any approach based on the central role of lexis is in many ways more grammatical than any traditional grammar syllabus” (2000, p.8). From the corpus linguistics perspective, thus, a “lexically based grammar” should be emphasized in language teaching. The participants in this study held, and thus reinforced, this view. They believed that the most common expressions, of word combinations, are the most grammatical. This perception follows Halliday’s concept of lexis as the “most delicate grammar” (1991, p.42), that is, lexis and grammar are not two separate things, but interact in systematic ways.

Not surprisingly, most of the participants equated the corpus approach with learning collocation, because the corpus showed them the most common strings of words (i.e., collocation). They agreed that corpus use increased their awareness of the importance of collocation as well as the collocational knowledge of words. However, they attributed their enhanced collocational knowledge not just to the corpus, but to other
reasons, such as the writing course and their increased exposure to English texts. Still, corpus use contributed to heightening their awareness of collocation.

More importantly, through the corpus the students not only focused on forms, but also on meanings in seeking out collocation patterns. They recognized that patterns using the same words could differ according to the nuances in surrounding texts. As Jabbour (2001) notes, the corpus approach generally focuses on both forms and functions. That is, the corpus approach is not only a mechanical process of looking for forms, but is also concerned with meanings and subtle nuances.

Increased language sensitivity, or language awareness, was another important influence of corpus use on participants’ language learning. Developing an awareness that collocations exist or that they are important in language learning/writing was an educational process in itself. Once participants were aware of the importance of collocations or common usage patterns, they were sensitized to those patterns, and most of them reported that they began to pay more attention to the combinations of words in their reading as well as their writing. In other words, corpus use not only helped them to solve immediate writing/language problems, but also expanded their language awareness and helped them with their language learning. In this regard, Tribble (2002) proposes that one of the most salient contributions of corpus analysis is “contextual and linguistic awareness raising during an EAP course” (p.133), and the results of this study lend strength to that assertion.

That the corpus approach heightened the students’ language awareness has important ramifications for language learning and teaching. As some researchers have
argued, better descriptions of language produced by corpus linguistics may not necessarily improve language teaching (Widdowson, 1991). Cook (1998) and Widdowson (1991) claim that the mere collection of native speakers’ uses of the language is not necessarily beneficial to students in different contexts. On the contrary, Widdowson (1996) argues, “authentic language is, in principle, incompatible with autonomous language teaching” (p.68). This study’s findings, however, suggest that the corpus approach has, at least, a secondary (if not a direct) effect on language learning. The corpus approach promoted the students’ self-monitoring in their L2 learning and writing. This suggests that awareness of the importance of word usage and a focus on lexis can add meaningful contexts to the learning process. While a question still remains as to when and how such awareness translates into actual learning, conscious language sensitivity may possibly promote learning and thus has its own importance and role in language learning and teaching. As such, it may be advisable to integrate corpus resources into conventional language teaching. The use of authentic language is not problematic in itself: what counts is how we approach the authentic texts and make the best use of them. When students need actual language input for more competent communication and L2 learning, as McCarthy (2001) points out, the concerns raised by Cook and Widdowson, among others, may be too negative with respect to the potential benefits of corpora by only drawing upon a ‘corpus-driven’ approach and failing to explore an alternative ‘corpus-informed’ approach.

In this regard, this study lends support to Gavioli and Aston’s (2001) view of a link between learner autonomy and authenticity. They assert that the learner, as a
discourse participant as well as an observer of real texts, constructs a model while using corpus data and thereby authenticates the real texts in his/her own discourse rather than simply imitating them. In this study, the students used authentic texts attentively for their own purposes and experiences with corpora: they knew when and how to utilize the real texts to become more independent L2 writers. Thus, as Gavioli and Aston argue, authenticity and learner autonomy should be seen as complementary, not as contrary. Extending this point further, as Carter, Hughes, and McCarthy (1998) argue, while teachers and learners can decide whether to use the new materials (i.e., corpora), learners should not be “disempowered” by lack of access to such options.

The change in the participants’ understanding of the language, in turn, was reflected in their approaches to writing and the writing process. Regardless of their frequency of corpus use, their exposure to it made them aware of the importance of common usage and collocation in writing. They came to focus not only on content, but also on the linguistic domain. That is, the corpus served as a meaningful reference for language input and also served as a catalyst in helping them to become more attentive to their writing.

The participants’ overall writing process did not dramatically change after adopting the corpus. However, the general process underwent changes that were minor but useful. For example, they paid closer attention to word usage and collocation during writing and editing. They formed a habit of checking on their writing while composing, rather than writing only rough, unpolished drafts, and hence the base was built for independent learning. In other words, as the corpus approach was introduced and linked
to the writing process, they took more responsibility for their writing. The fact that corpora were available also gave them more confidence in the quality of their final product. As such, their overall confidence in writing increased, and they approached writing with less emotional stress. The fact that they had more control over their own language learning and writing is an important point given that the main writing difficulties they encountered later were in the language domain.

While the other participants reported that their confidence in L2 writing increased by the second quarter, June was an exception. She was not confident at all in L2 writing, mainly due to having fewer opportunities to write in her major field. Her overall inexperience in L1 academic writing was probably another factor. In fact, three participants, June, Sung, and Nick, were inexperienced in L1 academic writing. However, Sung was experienced in L2 academic writing in Japanese from his master’s studies, and Nick already had one year of experience in English academic writing prior to this study, leaving June as the only one who truly had little experience in academic writing in any language. This study assumed that the students were at the same level of writing proficiency. However, this study found, as in previous studies, that “second language students who may be placed in the same level of ESL instruction … may not be at the same stage developmentally in their writing and may have very different strengths and weaknesses which they bring to the academic reading-writing task” (Bosher, 1998, p.222).

Each participant had individual issues to overcome in L2 writing. In particular, June had various difficulties in writing, ranging from expressions and organization, to
idea generation, which can be interpreted as the characteristics of less proficient writers. The study confirmed Raimes’ (1985) conclusion that “what the less proficient writers need … is more of everything; more instruction and practice in generating, organizing, and revising ideas; more attention to the rhetorical options available to them; and more emphasis on editing for linguistic form and style. Attention to process is thus necessary but not sufficient” (p.250).

In this respect, a pedagogical concern arises here: should education be an individual activity or a truly social process? Do our methods encourage independent learning at the expense of collaboration? There is no doubt that promoting discovery learning and learner autonomy are important, especially given limits in instructional time. To some extent, the direct presence of a teacher is important to effective learning. In the final analysis, language learning is a process of socialization into the communicative patterns of another culture, and this requires opportunities for spontaneous communication between teachers and students and between students. Language should be used to communicate with other human beings rather than as an abstract academic exercise aimed at cognitive edification. This begs the question, “does the corpus approach discourage other beneficial aspects of language learning that require human interaction?” After all, language learning takes place primarily in the human environment, not within a machine. Ferris and Hedgcock’s (1998) conclusion in this context is illuminating:

Computers cannot and should not replace instruction by the teacher. Novice writers need instruction, modeling, and practice in various aspects of the composing process and of technical aspects of writing (e.g., citing sources,
Teaching and learning are a holistic process that cannot be easily replaced by mechanistic technological methods. How to instill self-discipline and personal responsibility in students is a bigger question, whose answers should be sought in adopting technology in instruction.

### 6.1.3 Collocation and L2 writing

The last research question explored the participants’ beliefs in the role of L1 in L2 collocation uses and the role of collocation in L2 writing. All but Nick believed that their native language’s linguistic patterns interfered with L2 collocation uses and writing, as each language carries unique collocation patterns. That finding is hardly surprising, except for the important fact that the participants carefully observed the differences between L1 and L2. Their experiences and learning gave them a sophisticated understanding of the differences between L1 and L2. In fact, their observations resembled the technical analyses of professional linguists. For example, June, in her own parlance, expressed the linguistic principle that lexicalized items with unambiguous referents can be correctly translated into other languages. In contrast, delexicalized notions with obscure (or several) meanings are particularly problematic in translating word usage, which is addressed in the literature (Jabbour, 1997). As users of more than one language, the participants were aware of contrastive linguistic and rhetoric features. Based on their pro forma linguistic observations, they applied their understanding of their
L1 to their L2 writing. They knew when to rely on and how to use their L1 as resources in L2 writing. It is also noteworthy that their learning helped them to discriminate between the different usage in L1 and L2 and how to avoid errors in those usage. This learning effect suggests an important point of focus for English language education.

The participants uniformly agreed that collocational knowledge was more crucial for second language production (i.e., speaking and writing) than for comprehension, as many researchers have noted (Bahns & Eldaw, 1993; Gabry_-Biskup, 1990; Howarth, 1996,1998; Nesselhauf, 2003). The students perceived that reading and listening did not require extensive collocational competence in understanding external content. In contrast, writing and speaking were perceived to demand proficient word combinations, given the nature of language production. In particular, they recognized the importance of collocation to successful writing for general purposes, and for formal writing. Further, the participants perceived collocation as the key factor in producing writing that approximates native production. That is, they recognized that a good command of collocation is characteristic of native-like writing. Or, as Wol put it, “collocation can be a key to differentiate natives from non-natives.” Likewise, they regarded correct uses of collocation as the epitome of advanced level English and writing proficiency. They considered that collocation is the benchmark for total fluency in L2 learning.

The participants believed that repeated reading and writing practice was the best way to improve collocation knowledge. Throughout the study, they reiterated the need for writing practice, which made them become aware of word combinations. Only through writing could they recognize their usage problems and learn the correct usage
patterns. They saw writing as the most effective tool for internalizing common usage patterns. Using collocation in writing increased their awareness of patterns, leading to deliberate or conscious learning. The process was synergetic: there was a mutual, reinforcing relationship between collocation knowledge and writing. They learned that collocational knowledge can improve writing, and at the same time, writing can enhance collocational familiarity by increasing awareness through exposure to language combination.

While the participants agreed on the importance of collocation in L2 learning and writing, they noted constraints in teaching collocation in EFL contexts. As EFL learners in their homelands, rarely exposed to English in everyday life, they differentiated between the pedagogical needs in teaching ESL and EFL. They believed that EFL contexts require teaching general rules first because natural language use is less possible in these contexts, while teaching collocation could enhance ESL students’ learning. What is noteworthy is that the participants recognized the value of their previous English language education, while they complained at the beginning of the study about its emphasis on grammar rules. The students perceived contextual differences between ESL and EFL situations and suggested different needs and teaching methods accordingly. This finding may challenge the techno-deterministic notion that new technology is superior to traditional practice. In essence, it demonstrates that traditional practice and new technology have a complementary effect on learning. Traditional models build the platform on which technological models of education can be launched. The point here is
to assess not whether technology replaces older methods, but rather how both can be combined to create the most effective learning synergy.

In contrast to the other participants who were cautious about integrating the corpus approach and teaching collocation in EFL situations, Xiaodong strongly believed in its need and potential benefits in his home country. He believed that Chinese students in China needed the corpus more than ESL learners in America because of the lack of resources for learning English in China. This is consistent with Conrad (2000), who argued that corpora can improve language learning in foreign countries, despite the lack of native speakers in the teaching context. The use of authentic materials can be more beneficial in EFL contexts where natural language use hardly takes place and no target language use is available outside the classroom. Unlike ESL students who are given opportunities to meet native speakers of the language and relearn the language rules outside class, EFL learners tend to heavily rely on learning materials to encounter the use of the language (Tomlinson, 1998). Likewise, classroom materials are often the foremost resources for language learning in EFL situations. Given the characteristics of EFL contexts, it appears to be essential that EFL classroom materials represent the target language in authentic use. In this way, then, corpora can be a more efficient resource for teaching foreign language skills. That is, the need for native speakers – often a rare resource in the foreign setting – can be minimized by using this technological resource. That is not to say that corpora can entirely replace human interaction. For example, language styles vary by age and gender, which are not distinguished in the corpora.
However, given the paucity of native speakers in the foreign context, corpora may be a good alternative to teaching descriptive linguistic skills.

### 6.1.4 Becoming L2 writers in EAP contexts

While students’ corpus use and writing experiences are indistinguishable in this study, this section summarizes major writing-related points and reviews the broader context of the students’ learning of L2 writing.

The individual students’ experiences in the ESL writing course were analyzed not only to understand their perspectives on the use of corpus technology, but also to illuminate their transitions from L1 writers into L2 writers. Throughout the study, one question recurred like a steady drumbeat: “What does it mean to become an L2 writer in an academic setting?” This is more than a writing question for ESL students, as they become engaged in different language and academic environments while learning about writing. In fact, one prominent finding from this study was that L2 writing became more than the participants’ process of producing a text or completing course assignments. Rather, it was a process of learning different writing conventions and participating in academic cultures and entering a discourse community as members of particular academic disciplines.

The participants’ previous English education seldom offered them opportunities to see writing as a process of disciplinary enculturation. They had learned English mainly by traditional grammar-translation methods. In particular, their L2 writing experiences were limited to producing short texts mainly by translation rather than through a genuine notion of composing in the L2. For most of them, the TWE was their
first serious experience in English composition. That is, they were rarely trained in
general L2 composition, let alone L2 academic writing. Academic writing, in most cases,
entails organizational and linguistic conventions that are well established in specific
academic disciplines. Accordingly, academic writing was a daunting challenge for the
newcomers in an American EAP setting due to not only having to address general
language issues in L2 writing, but also to learning about academic discourse
familiarization – learning different academic/writing conventions. This raises a
pedagogical question for L2 writing teachers about how to connect ESL writing
instruction to disciplinary writing. In this regard, Hirvela (1997) addressed two essential
objectives of EAP writing instruction: 1) to “provide generic discourse knowledge, such
as the general principles for the writing of a literature review,” and 2) to “create
opportunities for students to conduct their own explorations of the discourse they hope to
gain control of” (p.86).

L2 writing research has presented contrasting views on the roles of general
composition skills and L2 language proficiency in L2 writing. Process-oriented studies
emphasize writing strategies and expertise over linguistic aspects, assuming that writing
in L1 and L2 is not significantly different in nature (e.g., Zamel, 1982, 1983). In
contrast, other studies emphasize language aspects (and the product) over writing
expertise, arguing that L2 writing is inherently different from L1 writing, as language
issues are a major component of L2 composition (e.g., Raimes, 1985; Pennington & So,
1993). The findings of this study do not strongly support either side of the argument.
Instead, the participants held mixed views on the importance of writing expertise and
language proficiency in L2 writing, rather than emphasizing one over the other. The majority perceived that writing in L2 was not inherently different from writing in L1 in terms of idea generation and the writing process, which follows the process-oriented research position. But, at the same time, they recognized that L2 writing is unique in that it requires the use of a different language, as the product-oriented studies argued. They agreed that the linguistic domain was the most conspicuous area of difference between L1 writing and L2 writing. In short, the students viewed writing expertise and language skills as inseparable entities that serve and interact with each other to produce competent L2 writing. While it is highly plausible that the students’ advanced level writing and language proficiency contributed to this view, this study challenges the reductionist “either–or” approach that emphasizes one over the other in L2 writing instruction.

Whereas the participants held balanced views on the importance of general composition skills and language proficiency in L2 writing, they acknowledged that language skills take more time to become competent than composition skills. At the beginning, they reported L2 writing difficulties in four areas: linguistic aspects, organizational concerns, content knowledge, and difference in writing culture. Later, though, the students’ difficulties were mainly related to lexical accuracy or linguistic aspects. Only one participant, Jung, who was inexperienced even in L1 academic writing, still faced various difficulties from linguistic issues to idea generation to organization. In contrast, the other students felt much more confident about the other areas except for the linguistic domain. They attributed their success in the other areas mostly to the writing course and writing practice. In particular, the course was directed at
structural and organizational concerns to facilitate their overall writing process. As most of them had few experiences in L2 academic writing, they appreciated learning about the general conventions of academic writing and structuring academic papers. Their enhanced content knowledge and participation in the native language environment also made their L2 writing easier. In addition, all students, except Jung, were already experienced in academic writing in other languages (mostly, in their native language). This helped them with the acquisition of composition skills in L2 writing.

Although the students acquired organizational conventions of academic writing and became competent at developing ideas and structuring thoughts, they still struggled while seeking proper linguistic features to express their intended meanings. They acquired a certain level of writing expertise and strategies in L2 writing within a comparatively short time through instruction and practice. However, their linguistic concerns could not be resolved in such a short time, and language issues lingered. That is, the linguistic domain remained a major challenge even for the advanced level students, who were well trained in composition skills. In this respect, this study’s findings confirm Raimes’ (1985) conclusion that “students should be taught not only heuristic devices to focus on meaning, but also heuristic devices to focus on rhetorical and linguistic features after the ideas have found some form” (pp.247-248). Whether this is only true for advanced level students with extensive experience in L1 academic writing, like in this study, or whether this also applies to beginning and intermediate L2 learners is a topic worthy of further exploration.
Throughout the study, the participants reiterated their difficulty in acquiring word usage and lexical accuracy as L2 learners. They believed that with practice they could become competent at organization and structure, but they felt they needed much more training to advance their writing skills in the language domain. The finding suggests that the lexical and language issues seem to be long-lasting problems even for advanced L2 writers who have prior knowledge and L1 writing experiences. This finding lends support to Hinkel’s (2002) call for directing greater attention to the linguistic and rhetorical features in L2 writing instruction. She argued that advanced L2 writers had “little trouble producing text according to the norms of the process-centered teaching of writing, to which most are exposed for a number of years during their intensive and academic writing instruction” (p.54). Instead, she emphasizes the need for addressing “the issue of language development, lexicon expansion, or sentence construction for NNSs [non-native speakers]” (p.56) in writing instruction. For L2 writers, the focus on content without taking into account their linguistic concerns may not lead them to become competent writers. The students need to be taught about the linguistic and rhetorical features specific to composing in L2.

The study found that the students were aware of their own weaknesses in L2 writing and engaged in self-help activities to overcome their writing difficulties. Many participants brought rich experiences in writing in their native language to the new context of L2 writing. They even took the initiative to construct a text bank that served as a small corpus of material to draw from. That students took ownership of the learning process was an important finding of the study. Whether this was a function of corpus use
was unclear. But this ownership model was prevalent in their learning. In other words, rather than passively receiving “knowledge,” students integrated it into their personal schema built on their own needs and experiences.

To conclude, the participants’ self-perceptions as L2 learners revealed how they used their previous learning/writing experiences as resources to adjust to the different writing culture, struggles, and difficulties they encountered. Along the way, they came to identify themselves as members of a discourse community as well as L2 writers whose knowledge of foreign writing conventions is required for academic success.

6.2 Pedagogical implications

The findings of the study provide significant implications for teaching L2 writing in an EAP classroom, especially where corpus use is integrated into instruction.

This study found that most participants had few experiences in English academic writing before they started graduate studies in America. While they were taught to produce short texts mainly by translations, the TWE was their first real exposure to English composition. If their experiences are typical of other EFL students, then much needs to be done to revamp the educational strategy that slights writing skills in EFL curricula. The educational system is certainly shortchanging the students’ L2 writing learning. A new model for L2 writing teaching is required that “sees writing not just as one of the language skills to be learned, or the last skill to be learned, but as an effective way for a learner to generate words, sentences, and chunks of discourse and to communicate them in the new language” (Raimes 1985, p.252). It is hoped that the
results of this study will help redress these shortcomings in order to develop effective writing pedagogies that view composing texts as an integral part of L2 learning.

One important point emerging from the study is that the linguistic features require advanced, trained L2 learners more time and experience to grasp due to their subtleties than is often assumed. The participants’ main difficulties that occurred later involved properly employing linguistic features of academic writing. While recent writing research mainly addresses the global or discoursal aspects of writing, language issues should not be ignored. For L2 writers, their mastery of lexical and grammatical accuracy can then contribute to increased confidence as a writer as well as about writing itself. This suggests a need for a pedagogy that accounts for a longer-term language learning process to acquire expressions and word usage than is recognized in conventional class preparation.

L2 writing entails double missions: within and across disciplinary needs. It is a holistic process that crosses disciplinary boundaries as well as situates itself in the bounds of specific disciplines. In other words, L2 writing requires providing disciplinary-specific knowledge about writing conventions as well as general features of writing, such as linguistic choices. How to capture both aspects within a pedagogical schema is a challenge that confronts EAP writing teachers. However, it is also imperative that content course faculty should cooperate with EAP writing teachers to address this issue, and vice versa. As newcomers, most participants in this study wanted feedback from content teachers not only on disciplinary content, but also on writing. This urgently calls for collaboration across disciplinary boundaries.
With respect to corpus use in L2 writing, the participants were favorable overall to its usefulness, despite their various difficulties in using the corpus. The ultimate question, then, is how can we harness the strengths and ameliorate the shortcomings of corpora as a viable tool for L2 education? One important benefit of the corpus approach that was confirmed in this study is its focus on collocation patterns and typical contexts of word uses. The focus on these commonly used chunks can lead L2 learners to acquire command of conventional use and fluency, which is often not achieved by studying structural rules. Thus, the study suggests a strategy that integrates lexico-grammatical features into language education and writing instruction. The combined concept of lexis and grammar can enhance students’ repertoire of English expressions and usage in L2 writing.

A significant finding of the study is that students took more responsibility for their language learning as a result of their corpus experience. They became more independent writers and increased their confidence in L2 writing through increased access to linguistic resources. This is one of the most important roles that corpus technology plays in L2 writing. Corpora are a tool that allows students to independently solve their linguistic and writing problems, and it raises students’ awareness through the use of authentic texts. As such, L2 writing instructors can integrate the student initiated corpus search into their teaching (rather than replace the conventional ways) in order to promote students’ independence and self-monitoring in L2 writing. Applying the corpus approach to L2 writing instruction can enrich the instructional environment by heightening the students’ responsibility in the learning process.
In this study, the course instructor’s adept use of general corpora and discipline-specific class activities provide a valuable model for EAP writing teachers. In addition to the general corpora that were used to supply linguistic input to students, the instructor offered complementary, alternative materials and activities, e.g., constructing a style manual for the common structure of papers specific to a given field based on searches for academic journal articles. This can be seen as a specialized, miniaturized type of corpus construction. This kind of class activity can cater to the specific needs of the students from various academic fields. At the same time, this approach can be adopted and expanded in EAP writing instruction: to construct specialized corpora to include a broader range of expressions and typical language found in specific academic disciplines, thus showing discourse-level characteristics of disciplinary writing beyond the sentence-level text.

The study found the students to be careful observers of their L1 and L2. They brought a rich and sophisticated foundation of multilingual knowledge to the L2 writing experiences. Teachers can tap into this knowledge to create more effective pedagogical practice rather than merely leaving students to their own personal learning devices for L2 learning. Likewise, it is important for L2 educators to understand the effect of L1 on L2 learning and writing. The question is not whether the L1 interferes, but when it can be a positive resource and when it hinders language learning. For instance, the study found that differences in L1 and L2 usage do not always translate into errors when students are aware of the differences through their learning. This suggests the usefulness of contrastive analysis that identifies the conspicuous differences between languages.
Through the use of such analysis, the students’ previous experiences with certain languages can become more valuable in L2 learning and writing.

In the present study, a wide variety of individual experiences and learning contexts were involved in deciding the levels of the students’ willingness and success in using corpora. Writing teachers who attempt to incorporate the corpus approach into their instruction need to understand the multi-faceted aspects of technology use that could facilitate or impede the individual students’ L2 writing. It is clear that not all students learn in the same way, and the new technology may not work as well for all students. In particular, one important pedagogical question to be addressed is how to meet the needs of the students who are at variant levels in their development of academic literacy, including those in the same course, who are presumed to have about the same amount of L2 writing ability. For example, various writing difficulties of less proficient students will probably interfere with their corpus use. L2 educators need to identify individual difficulties and needs in students’ L2 writing and help the students build upon their strengths so as to develop skills that overcome their weaknesses.

With respect to collocation learning, this study’s findings indicate that collocation can be easily applied to any level of L2 learning. This study suggests that although lower-level learners may have difficulty manipulating corpus data, they can still benefit from exposure to collocational aspects through other sources, such as printed materials. It is also important to note that corpus findings can be used beneficially to develop teaching materials, with due attention paid to the students’ levels of language and writing.
proficiency. For instance, corpus-informed materials can enrich the learning process as a handout designed to complement textbook materials provided to the students.

Lastly, serious thought should be given to ways to manage unexpected technology malfunctions that could occur during the course of instruction, a frustration that may inhibit the learning of some students, as seen in this study when some students were left helpless after the online corpus broke down. They did not even seek help from alternative resources. It is imperative that students have comparable services available so they can continue studies in the event of technological breakdown. Other alternatives include corpora available off-line, such as on a CD-ROM, and dictionaries that are created via corpora, such as the Collins Cobuild dictionaries. In fact, corpus-based dictionaries are now becoming more available due to the development of computer technology and the enhanced recognition of the usefulness of corpora. Although such dictionaries do not allow students to use corpus resources to the fullest, they can at least provide them with the most common contexts of words, which is preferable to giving them out-of-context definitions.

6.3 Limitations of the study

One limitation of this study was its duration, which prevented observation of long range, meaningful changes in students’ L2 writing approaches and processes. As Bernhardt et al. (1989) put it, “the effects of computers on writing ability may not be a matter of quick transfer, but of subtle and incremental evolution over the life of a writer” (p.129). Students likely need more extended exposure to the new technology before its results become apparent in their actual writing.
Another limitation is that the participants were exposed only to the use of general corpora rather than to specialized corpora. The corpus use in this study was aimed at helping the students deal with their immediate linguistic problems, and its value was confirmed in the findings, but it would also be interesting to explore how students use specialized corpora in their EAP writing. The field of EAP/ESP is often considered one of the main beneficiaries of corpus analysis because courses in this domain can be linked to specific disciplinary needs. It would be worthwhile to examine whether students exhibit the same purposes and patterns of, as well as attitudes toward, corpus use while using discipline-specific corpora.

This study primarily investigated students’ writing experiences rather than examining their writing products in relation to corpus use. A limitation of this study, then, is the lack of focus on written product. Any conclusion cannot be drawn from this study about the direct relationship between corpus use and writing development. A valuable future line of inquiry would examine whether corpus use actually effects changes in the students’ writing proficiency.

Lastly, we cannot conclusively generalize from a small sample consisting of a limited number of participants. However, the qualitative approach helps us understand processes of learning that are not easily identifiable through quantitative methods. Some of the insights gained can be further examined in quantitative studies that investigate their applicability over a broader range and larger number of participants.
6.4 Recommendations for future research

The present study provided a deeper understanding of corpus practices in L2 writing (instruction). This study harvested many insights into L2 writing in a specific context with a limited number of participants. The study discovered the benefits of corpus use in the specific context of advanced ESL students’ L2 writing at a large midwestern American university. Further studies are needed to expand upon these insights in order to more fully understand corpus pedagogy. Teaching is extremely sensitive to a given context. It is hard to apply a universal principle across all contexts, so the evaluation of the corpus approach in a variety of contexts is greatly needed, especially when the range of technological resources available for pedagogical use continues to expand. We need to do more research to identify the most effective teaching strategies to fulfill this critical need. For instance, this study proposed the potential benefits of corpus use in foreign language teaching contexts, so future research can examine EFL students’ corpus use and the advantages and caveats of corpus pedagogy in the foreign language context. The important question is how to adapt the corpus approach successfully to a particular learning context to enhance students’ academic literacy experiences.

The findings point also to the need for further investigations into uses of different kinds of corpora and their impact on L2 writing processes. This study examined general corpora with a sentence-focused data presentation, while the instructor encouraged students to manually compile an alternative small corpus of language uses found in their own fields. It would be worthwhile to expose the students to longer, well-designed
corpora specific to their own disciplines and see how they use them. Considering that many researchers argue for the usefulness of specialized corpora in EAP disciplinary writing, it is imperative that the next phase of research should study specialized corpora that are created for specific academic disciplines.

This study did not analyze students’ writing products to explore any changes in their writing performance. A fruitful line of inquiry would be to examine their written products to analyze improvements made. Also, this study presented the students’ perceptions of the role of their native language in L2 collocational uses and writing. Further research should examine L2 collocational uses in students’ writing to see whether their textual practices are consistent with their perceptions. Another line of inquiry could examine L2 learners’ most common collocation errors and conduct contrastive analyses in linguistic patterns between languages. Such a study would be invaluable for future teachers.

Lastly, one important future research area is to apply corpus instruction to different varieties of English. Although the online corpora used in this study are based on Standard American and British English, English is continually transformed into various dialects and subvarieties around the world, and many Englishes exist. In fact, the issue of representativeness is of a great concern in the production and use of corpora. With the growing recognition of their importance and the enormous advancement of computer capacity, a variety of corpora (e.g., international corpora, learner corpora, and multilingual corpora) are now being produced. The pedagogical selection of corpora in terms of varieties is an important point if corpora are to be used in regional or even global
contexts. This touches on the important (and ethical) question, whose English are students learning?


APPENDIX A

RECRUITING SURVEY

Each session of the class will be observed for research purposes regarding using the corpus in ESL writing courses.

Classroom observation will enable us to document classroom activities in order to:

- Explore how the corpus is used in ESL writing instruction
- Examine how ESL students respond to its use and how it helps them enhance their writing skills
- Analyze classroom interaction/dynamics regarding its use
- Suggest pedagogical implications for using corpus technology in ESL writing curriculum

I understand that this class will be observed for research purposes.

______________________________________________________
(Print Name) Signature Date

The following asks about your information, which may be used for screening participants.

1. Name __________________________ 2. Email address __________________________
3. Gender: Male ______ Female ______ 4. Program: Master ______ Ph.D. ______
5. Major subject at OSU (e.g., Biology) __________________________
6. Country you are from _________________ 7. Native language __________________________
8. How long have you been in America? _______Years ________Months
9. How long have you studied English? About ______Years
10. In general, do you like to write in English? Yes ______ No ______
11. Have you taken any other ESL writing courses before this quarter? If so, what were they?
12. How many academic papers do you need to write this quarter for all your courses and research? (e.g. essays, term-papers, or research papers) __________________________
13. In general, do you like to use computer? Yes ______ No ______
14. Do you have Internet access at home? Yes ______ No ______
15. Have you used a corpus (e.g., Collins Cobuild corpus) before this class? Yes____ No____

I am interested in participating in this research. Yes _____ No _____ Not decided _____
APPENDIX B

STYLE ANALYSIS SURVEY (SAS):
Assessing Your Own Learning And Working Styles

Purpose:
The SAS is designed to assess your general approach to learning and working. It does not predict your behavior in every instance, but it is a clear indication of your overall style preferences.

Instructions:
For each item circle the response that represents your approach. Complete all items. There are five major activities representing five different aspects of your learning and working styles. At the end you will find a self-scoring key and an interpretation of the results.

Timing:
It generally takes about 30 minutes to complete the SAS. Do not spend too much time on any item. Indicate your immediate response and move on to the next item.

For each item, circle your immediate response:
0 = Never, 1 = Sometimes, 2 = Very Often, 3 = Always

Activity 1: How I Use My Physical Senses to Study or Work

<table>
<thead>
<tr>
<th>Number</th>
<th>Statement</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I remember something better if I write it down.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>I take lots of notes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I can visualize pictures, numbers, or words in my head.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>I prefer to learn with video or TV more than with other media.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>I underline or highlight the important parts I read.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>I use color-coding to help me as I learn or work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>I need written directions for tasks.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>I get distracted by background noises.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>I have to look at people to understand what they say.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>I am more comfortable when the walls where I study or work have posters and pictures.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
11. I remember things better if I discuss them out loud.  
12. I prefer to learn by listening to a lecture or a tape, rather than by reading.  
13. I need oral directions for tasks.  
14. Background sounds help me think.  
15. I like to listen to music when I study or work.  
16. I can easily understand what people say even if I can’t see them.  
17. I remember better what people say than what they look like.  
18. I easily remember jokes I hear.  
19. I can identify people by their voices.  
20. When I turn on the TV, I listen to the sound more than watching the screen.  

21. I’d rather just start doing things rather than pay attention to directions.  
22. I need frequent breaks when I work or study.  
23. I move my lips when I read silently.  
24. I avoid sitting at a desk when I don’t have to.  
25. I get nervous when I sit too long.  
26. I think better when I can move around.  
27. Manipulating objects helps me to remember.  
28. I enjoy building or making things.  
29. I like a lot of physical activities.  
30. I enjoy collecting cards, stamps, coins, or other things.  

For each item, circle your immediate response:  
0 = Never, 1 = Sometimes, 2 = Very Often, 3 = Always

**Activity 2: How I Deal with Other People**

1. I prefer to work or study with others.  
2. I make new friends easily.  
3. I like to be in groups of people.  
4. It is easy for me to talk to strangers.  
5. I keep up with personal news about other people.  
6. I like to stay late at parties.  
7. Interactions with new people give me energy.  
8. I remember people’s names easily.  
9. I have many friends and acquaintances.  
10. Whenever I go, I develop personal contacts.  
11. I prefer to work or study alone.  
12. I am rather shy.  
13. I prefer individual hobbies and sports.  
14. It is hard for most people to get to know me.  
15. People view me as more detached than sociable.  
16. In a large group, I tend to keep silent.  
17. Gatherings with lots of people tend to stress me.  
18. I get nervous when dealing with new people.
19. I avoid parties if I can. 1 2 3 4
20. Remembering names is difficult for me. 1 2 3 4

For each item, circle your immediate response:
0 = Never, 1 = Sometimes, 2 = Very Often, 3 = Always

**Activity 3: How I Handle Possibilities**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I have a vivid imagination.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>I like to think of lots of new ideas.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>I can think of many different solutions to a problem.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>I like multiple possibilities and options.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>I enjoy considering the future events.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>Following a step-by-step procedure bores me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>I like to discover things rather than have everything explained.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>I consider myself original.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>I am an ingenious person.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>It feels fine if the teacher or boss changes the plan.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>I am proud of being practical.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>I behave in a down-to-earth way.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>I am attracted to sensible people.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>I prefer realism instead of new, untested ideas.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>I prefer things presented in a step-by-step way.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>I want a class or work session to follow a clear plan.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>I like concrete facts, not speculation.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>Finding hidden meanings is frustrating or irrelevant to me.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td>I prefer to avoid too many options.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>I feel it is useless for me to think about the future.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

For each item, circle your immediate response:
0 = Never, 1 = Sometimes, 2 = Very Often, 3 = Always

**Activity 4: How I Approach Tasks**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I reach decisions quickly.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>I am an organized person.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>I make lists of things I need to do.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>I consult my lists in order to get things done.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Messy, unorganized environments make me nervous.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>I start tasks on time or early.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>I get places on time.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Deadlines help me organize work.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>I enjoy a sense of structure.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>I follow through with what I have planned.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
11. I am a spontaneous person.  
12. I like to just let things happen, not plan them.  
13. I feel uncomfortable with a lot of structure.  
14. I put off decisions as long as I can.  
15. I have a messy desk or room.  
16. I believe deadlines are artificial or useless.  
17. I keep an open mind about things.  
18. I believe that enjoying myself is the most important thing.  
19. Lists of tasks make me feel tired or upset.  
20. I feel fine about changing my mind.

For each item, circle your immediate response:  
0 = Never, 1 = Sometimes, 2 = Very Often, 3 = Always

**Activity 5: How I Deal with Ideas**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I prefer simple answers rather than a lot of explanations.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Too many details tend to confuse me.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>I ignore details that do not seem relevant.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>It is easy for me to see the overall plan or big picture.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>I can summarize information rather easily.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>It is easy for me to paraphrase what other people say.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>I see the main point very quickly.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>I am satisfied with knowing the major ideas without the details.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>I can pull together (synthesize) things easily.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>When I make an outline, I write down only the key points.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>I prefer detailed answers instead of short answers.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>It is difficult for me to summarize detailed information.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>I focus on specific facts or information.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>I enjoy breaking general ideas down into smaller pieces.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>I prefer looking for differences rather than similarities.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>I use logical analysis to solve problems.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>My written outlines contain many details.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>I become nervous when only the main ideas are presented.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>I focus on the details rather the big picture.</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>When I tell a story or explain something, it takes a long time.</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
APPENDIX C

STRATEGY INVENTORY FOR LANGUAGE LEARNING (SILL)
Version for Speakers of Other Languages Learning English

Version 7.0 (ESL/EFL)

Directions

This form of the STRATEGY INVENTORY FOR LANGUAGE LEARNING (SILL) is for students of English as a second or foreign language. You will find statements about learning English. Please read each statement. On the separate Worksheet, write the response (1, 2, 3, 4, or 5) that tells HOW TRUE OF YOU THE STATEMENT IS.

1. Never or almost never true of me
2. Usually not true of me
3. Somewhat true of me
4. Usually true of me
5. Always or almost always true of me

NEVER OR ALMOST NEVER TRUE OF ME means that the statement is very rarely true of you.

USUALLY NOT TRUE OF ME means that the statement is true less than half the time.

SOMEWHAFT TRUE OF ME means that the statement is true of you about half the time.

USUALLY TRUE OF ME means that the statement is true more than half the time.

ALWAYS OR ALMOST ALWAYS TRUE OF ME means that the statement is true of you almost always.

Answer in terms of how well the statement describes you. Do not answer how you think you should be, or what other people do. There are no right or wrong answers to these statements. Put your answers on the separate Worksheet. Please make no marks on the
items. Work as quickly as you can without being careless. This usually takes about 20-30 minutes to complete. If you have any questions, let the teacher know immediately.

1. Never or almost never true of me
2. Usually not true of me
3. Somewhat true of me
4. Usually true of me
5. Always or almost always true of me

Part A
__1. I think of relationships between what I already know and new things I learn in English.
__2. I use new English words in a sentence so I can remember them.
__3. I connect the sound of a new English word and an image or picture of the word to help me remember the word.
__4. I remember a new English word by making a mental picture of a situation in which the word might be used.
__5. I use rhymes to remember new English words.
__6. I use flashcards to remember new English words.
__7. I physically act out new English words.
__8. I review English lessons often.
__9. I remember new English words or phrases by remembering their location on the page, on the board, or on a street sign.

Part B
__10. I say or write new English words several times.
__11. I try to talk like native English speakers.
__12. I practice the sounds of English.
__13. I use the English words I know in different ways.
__15. I watch English language TV shows spoken in English or go to movies spoken in English.
__16. I read for pleasure in English.
__17. I write notes, messages, letters, or reports in English.
__18. I first skim an English passage (read over the passage quickly) then go back and read carefully.
__19. I look for words in my own language that are similar to new words in English.
__20. I try to find patterns in English.
__21. I find the meaning of an English word by dividing it into parts that I understand.
__22. I try not to translate word-for-word.
__23. I make summaries of information that I hear or read in English.
Part C
___24. To understand unfamiliar English words, I make guesses.
___25. When I can’t think of a word during a conversation in English, I use gestures.
___26. I make up new words if I do not know the right ones in English.
___27. I read English without looking up every new word.
___28. I try to guess what the other person will say next in English.
___29. If I can’t think of an English word, I use a word or phrase that means the same thing.

Part D
___30. I try to find as many ways as I can to use my English.
___31. I notice my English mistakes and use that information to help me do better.
___32. I pay attention when someone is speaking English.
___33. I try to find out how to be a better learner of English.
___34. I plan my schedule so I will have enough time to study English.
___35. I look for people I can talk to in English.
___36. I look for opportunities to read as much as possible in English.
___37. I have clear goals for improving my English skills.
___38. I think about my progress in learning English.

Part E
___39. I try to relax whenever I feel afraid of using English.
___40. I encourage myself to speak English even when I am afraid of making a mistake.
___41. I give myself a reword or treat when I do well in English.
___42. I notice if I am tense or nervous when I am studying or using English.
___43. I write down my feelings in a language learning diary.
___44. I talk to someone else about how I feel when I am learning English.

Part F
___45. If I do not understand something in English, I ask the other person to slow down or say it again.
___46. I ask English speakers to correct me when I talk.
___47. I practice English with other students.
___48. I ask for help from English speakers.
___49. I ask questions in English.
___50. I try to learn about the culture of English speakers.
APPENDIX D

CORPUS SEARCH LOG FORM

Name: ______________________________

<table>
<thead>
<tr>
<th>Date</th>
<th>Search items</th>
<th>For what purpose?</th>
<th>Helpful?</th>
<th>Time spent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E

SURVEY QUESTIONNAIRE

INSTRUCTIONS

Thank you for your participation in this study. This questionnaire asks you the questions concerning your experiences in using Collins Cobuild Corpus. This is not a test so there are no “right” or “wrong” answers. Your sincere answers would be most appreciated.

Computer Use and Experience

This section asks about your computer experience. Please mark/write your answer.

1. How long have you used a computer? About _____ Years

2. Generally speaking, how many hours a day do you use a computer for school work?

   0       1       2          3            4              5              6      7        8              9        10 hours

3. Generally speaking, how many hours a day do you use a computer for personal purposes?

   0       1       2          3            4              5              6      7        8              9        10 hours

4. When you use a computer for personal purposes, do you use English or your native language?

   ___ English
   ___ Native language
   ___ Both
   ___ Others (specify):

5. How much of your total computer time is in your native language?

   0%     10%      20%        30%        40%        50%        60%     70%       80%        90%      100%
6. Do you have Internet access at home?
   ____ Yes
   ____ No

7. If yes, do you use the Internet through telephone modem or Cable modem/LAN at home?
   ____ Telephone modem
   ____ Cable modem/LAN

**Dictionary and Corpus Use**

This section asks about your dictionary and corpus use. Please mark/write your answer.

8. Do you use a dictionary for English writing?
   ____ Yes
   ____ No

9. If yes, what kind of dictionary do you often use? (Check all the numbers that apply)

<table>
<thead>
<tr>
<th>Language</th>
<th>Type</th>
<th>Paper</th>
<th>On-line</th>
<th>Computer program</th>
<th>Electronic handheld</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual (e.g., English-English)</td>
<td></td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Bilingual (e.g., Chinese-English)</td>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

10. Had you heard about corpora before you took this class?
    ____ Yes
    ____ No

11. Had you used any corpus before you took this class?
    ____ Yes (if yes, which one? ______________________________________)
    ____ No

12. Do you generally use Collins Cobuild Corpus for your composing and/or revising papers?
    ____ Composing
    ____ Revising
    ____ Both
    ____ Others (specify):

13. Do you generally use the corpus at home and/or school?
    ____ Home
    ____ School
    ____ Both
14. List any reasons that you do not use the corpus more than you do now. If there is more than one reason, please number them using (1) as the strongest reason.

____ I prefer a dictionary.
____ It takes too much time for me.
____ I am usually not connected to the Internet when I’m writing.
____ I don’t have time.
____ I don’t trust the corpus data.
____ I am not interested.
____ It is not useful for me. (if checked, why?: ____________________________)
____ Others (specify):

Advantages of Corpus Use

This section asks about your perceptions of advantages of corpus use in your English writing. Please rate the following, on a scale from 1 to 6, according to how much you agree with the statements by circling the appropriate number.

1=strongly disagree; 2=disagree; 3=slightly disagree; 4=slightly agree; 5=agree; 6=strongly agree

1. The corpus is more helpful than a dictionary for English writing.............. 1 2 3 4 5 6
2. Using the corpus is helpful for learning the meaning of vocabulary....... 1 2 3 4 5 6
3. Using the corpus is helpful for learning the usage of vocabulary........... 1 2 3 4 5 6
4. Using the corpus is helpful for learning the usage of phrases.............. 1 2 3 4 5 6
5. Using the corpus is helpful for learning grammar............................. 1 2 3 4 5 6
6. Using the corpus is helpful for improving my English reading skill..... 1 2 3 4 5 6
7. Using the corpus is helpful for improving my English writing skill..... 1 2 3 4 5 6
8. Using the corpus is helpful for improving my English academic writing.. 1 2 3 4 5 6
9. Using the corpus is helpful for increasing my confidence about English writing................................................................. 1 2 3 4 5 6

Difficulties/Problems in Corpus Use

This section asks about your perceptions of difficulties in corpus use in your English writing. Please rate the following, on a scale from 1 to 6, according to how much you agree with the statements by circling the appropriate number.

1=strongly disagree; 2=disagree; 3=slightly disagree; 4=slightly agree; 5=agree; 6=strongly agree

10. I have some difficulty in using the corpus due to limited access to computer/Internet................................................................. 1 2 3 4 5 6
11. I have some difficulty in using the corpus due to low speed of Internet connection ................................................................. 1 2 3 4 5 6
12. I have some difficulty in using the corpus due to unfamiliar vocabulary on concordance output (40 sentences) or collocate output (frequency data) .................................................................................. 1 2 3 4 5 6
13. I have some difficulty in using the corpus due to cut-off sentences in concordance output................................................................. 1 2 3 4 5 6
14. I have some difficulty in using the corpus due to too few sentences in concordance output ........................................................................ 1 2 3 4 5 6
15. I have some difficulty in using the corpus due to too many sentences in concordance output ............................................................... 1 2 3 4 5 6
16. I have some difficulty in analyzing concordance output................... 1 2 3 4 5 6
17. I have some difficulty in analyzing collocate output ....................... 1 2 3 4 5 6
18. I have some difficulty in using the corpus because data analysis takes too much time for me ............................................................. 1 2 3 4 5 6
19. I have some difficulty in using the search technique ...................... 1 2 3 4 5 6
20. The real texts of the corpus are too difficult to understand .............. 1 2 3 4 5 6
21. The corpus is not helpful for my academic writing because it does not apply to my field ................................................................. 1 2 3 4 5 6

**Corpus Use in English Writing**

This section asks about your reactions to using corpus technology in your English writing. Please rate the following, on a scale from 1 to 6, according to how much you agree with the statements by circling the appropriate number.

1=strongly disagree; 2=disagree; 3=slightly disagree; 4=slightly agree; 5=agree; 6=strongly agree

22. I feel comfortable using corpus technology ....................................... 1 2 3 4 5 6
23. I often use the corpus for my English writing .................................... 1 2 3 4 5 6
24. When I have problems with English writing, I search for help in the corpus ....................................................................................... 1 2 3 4 5 6
25. When I search for information in the corpus, I usually get the information that I need .............................................................. 1 2 3 4 5 6
26. The corpus is a useful resource for my English writing ................... 1 2 3 4 5 6
27. Using the corpus has changed my understanding of the language ....... 1 2 3 4 5 6
28. Using the corpus has changed my approach to English writing ...... 1 2 3 4 5 6
29. Using the corpus has changed the process of my English writing ... 1 2 3 4 5 6
30. As I have used the corpus more, I have come to like it more ............ 1 2 3 4 5 6
31. If I had used the corpus earlier, I would have had a better score on TOEFL .............................................................. 1 2 3 4 5 6
Corpus Use in Writing Instruction

This section asks about your reactions to using corpus technology in ESL writing instruction. Please rate the following, on a scale from 1 to 6, according to how much you agree with the statements by circling the appropriate number.

1 = strongly disagree; 2 = disagree; 3 = slightly disagree; 4 = slightly agree; 5 = agree; 6 = strongly agree

32. It is easy to construct prototypes by use of concordance/collocate output………………………………………………………… 1 2 3 4 5 6
33. I feel confident in constructing prototype strings………………….. 1 2 3 4 5 6
34. Constructing prototypes is a useful experience for language learning………………………………………………………….. 1 2 3 4 5 6
35. I use the prototype strings the teacher gave us as a reference in my writing…………………………………………………………. 1 2 3 4 5 6
36. I often use the corpus by my own choice…………………………… 1 2 3 4 5 6
37. The corpus was helpful for composing my papers for this course… 1 2 3 4 5 6
38. The corpus was helpful for revising my papers for this course…… 1 2 3 4 5 6
39. I use the corpus when writing papers for other courses too……….. 1 2 3 4 5 6
40. I will use the corpus for my English writing in the future…………. 1 2 3 4 5 6
41. I recommend using the corpus to other international students …… 1 2 3 4 5 6
42. I recommend using the corpus in 108.02 in future quarters………. 1 2 3 4 5 6
43. The corpus should be introduced in all ESL writing courses……… 1 2 3 4 5 6
44. The corpus should be introduced in English classes in my home country…………………………………………………………………. 1 2 3 4 5 6

Overall Evaluations of Corpus Use

This section asks about your overall evaluations of corpus use in the format of open-ended questions. Please write your answers in the blanks provided below.

45. What is the greatest advantage of using the corpus in your English writing?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

46. What is the greatest disadvantage of using the corpus in your English writing?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Demographic Information

This section asks about your demographic information. Please mark/write your answer.

1. What is your gender?
   ____ Female
   ____ Male

2. What year were you born? ______________

3. What country are you from? ______________________________

4. What is your major? ______________________________

5. What degree are you currently seeking?
   ____ Master’s
   ____ Ph.D.

6. How long have you been in an American university setting? _____Years
   _____Months

7. What was your TOEFL score that you submitted for the application to this school?
   TOEFL score: _________
   Reading section score: _________
   Grammar section score: _________
   TWE score: _________
APPENDIX F

RECRUITMENT LETTER

Dear Student:

You are invited to participate in a research project entitled “An Investigation of Students’ Experiences with Corpus Technology in Second Language Academic Writing.” This research will be performed to complete a Doctor of Philosophy Degree in the College of Education. The data for this research will be collected during Autumn Quarter 2003.

This research aims to gain a deeper understanding of the role of corpus technology in second language (L2) academic writing. In particular, it aims to uncover the changes in ESL students’ writing experiences, associated with corpus use, and determine the extent to which the technology facilitates the writing process and contributes to L2 academic writing. Furthermore, the researcher hopes that this study will provide insights into using corpus technology in L2 writing instruction and suggest pedagogical implications for the teaching of L2 writing.

If you choose to participate in this study, you will be asked to give an interview once every two weeks, where the researcher will ask you questions about your experiences using corpus technology in your English writing. These interviews will be scheduled at your convenience and should not last more than an hour. The interviews will be audio-taped and transcribed. In addition, your writing drafts and corpus search logs for your ESL writing classes will be analyzed in order to identify any changes in your writing while using the corpus. Upon completion of the dissertation, the interview audiotapes will be erased, and the interview transcripts and your written work will be shredded. Also, in an attempt to preserve your confidentiality, you may choose a pseudonym at the beginning of the study. This pseudonym will be used throughout the data collection and the writing of the dissertation. Please note that those interviews and subsequent analysis of your work are specific to this research project and will have no impact on your course grades. Finally please also note that in appreciation of your participation in this study, the researcher would like to provide not only assistance in your writing of papers for other courses, but also book certificates.

It should be stressed that your participation is entirely voluntary. If at any time you decide that you do not want certain information to appear in the published record of this
research, you may say so to the researcher, who is bound to honor your request. During
the interviews, you will not have to answer any questions that make you uncomfortable.
More importantly, you will be able to withdraw from participation in this study at any
time without any penalty.

If you have any questions regarding this research or your rights related to participation in
this study, please feel free to contact the researcher. Your participation in this study
would be greatly appreciated. Thank you very much for your time and interest!

Best regards,

Hyunsook Yoon
APPENDIX G

INFORMED CONSENT FORM

Protocol # 2003EO199

CONSENT FOR PARTICIPATION IN RESEARCH

I consent to participating in research entitled: “An Investigation of Students’ Experiences with Corpus Technology in Second Language Academic Writing.”

Dr. Shelley Wong, Principal Investigator, or her authorized representative Hyunsook Yoon has explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described, as have alternative procedures, if such procedures are applicable and available.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study without prejudice to me.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date: _________________________         Signed: _______________________________
(Participant)

Signed: _______________________
(Principal Investigator or her authorized representative)

Signed: _______________________
(Person authorized to consent for participant, if required)

Witness: __________________________
APPENDIX H

EXCERPTS OF INTERVIEW TRANSCRIPTS TRANSLATED FROM KOREAN INTO ENGLISH

1. Wol

1) I don’t dislike writing in English because I was a Korean language arts teacher and I liked to write anyway. But I don’t like the fact that I can’t express myself fully in English as much as I can in Korean. I want to keep my own unique voice in English, but I can’t. … I can’t fully express what I am thinking because I can’t use a variety of vocabulary or sentence structure. I tend to use familiar expressions over and over again. So, all my sentences look very similar and they are not expanding.

2) Native-like writing 은 그 단어를 정확하게 정확한 용법을 알고 쓰는 거 하고. 좋은 writing 은 native writer 든 외국인 writer 든 간에 자기 생각을 잘 organize 해서 쓰는 거. 정확한 문법에 따라서 쓸 수는 있고 정확하게 잘 조작을 해서 쓸 수 있지만 non-native 는 오래된 속담이나 책을 인용한다든지 문구를 인용한다든지 잘 모르잡아요. 문구나 usage 나 표현을 외국인이 쓴 건지 알 수가 있죠. (11/14/03)

Native-like writing is one that employs accurate word usage. Non-natives can also become good writers if they use correct grammar and well organize their thoughts. But, we still can tell non-native writing from native writing in terms of usage and idiomatic expressions. They are different in those expressions.
2. Chan

1) 최근에는 실제로 많이 찾아 보는 게 내가 log 에 많이 적었는지는 모르겠지만, 형용사나 부사에요. 어떤 게 수식하는 게 가장 적합한가, “barrier”가 심각한 거는 뭐나, “severe”나, “heavy”나, “great”나 “major”나 뭐 이런 부분들이 난 한국에서 writing 가끔씩 할 때는 barrier 뭐 적당히 무겁고 심각하고 한 걸 대충 때려 쓰면은 별 문제가 없었는데, 이제는 그게 아 이게 뭐 다른 게 있겠구나 하는 생각을 많이 하고, 난 그 부분이 이번 writing 에서 가장 구체적으로 좀 바뀌는 터닝 하게 된 부분인 것 같아요. 이전에 사실은 동사나 명사나 전치사나 이 부분은 문장의 큰 골격이 수식하는 거는 뭐도 그만인데 그렇게 있어야지 문장이 좀 부드리워지고 풍부해지고 그런 걸 생각하니까 어색한 형용사 쓰기가 좀 어색하죠. (11/10/03)

Recently, what I look up the most in the corpus is actually adjectives and adverbs, that is, which words are the most appropriate to modify certain nouns or verbs. When I want to modify a “barrier,” is it “severe,” “heavy,” “great,” or “major”? I was not so much concerned about this aspect before. I just used any words. In fact, I thought that verbs, nouns and prepositions were important skeletons of sentences, but modifiers were trivial and they did not make a big difference. But, I came to realize that they also play an important role in the sentences. Those modifiers contribute to making much more fluent and eloquent sentences. So, now I am concerned about them a lot. I think this is the biggest change in my writing.

2) 큰 의미에서 글 쓰기 대한 approach 자체가 바꾼 건 아니지만, 글 쓰는 writing의 과정에서 이 부분이 들어오니까 절적으로나 과정적으로 바뀐 게 있죠. 과정적으로는 쉽게 내가 그냥 쓰다가 틀리던 뭐든 다 쓰고 프린트해서 다 다듬거나 하는 과정들이 전반적으로 바뀐 건 아니지만, 그 과정 사이에서 내가 문장이 안 끝났을 때 수시로 체크하는 곳이 있잖아요. Corpus를 확인하면서 이 표현이 어렇게 하면 가장 적합한가에 대해서 쓰는 과정에서 문장이 완전히 마무리 되지 않는 과정에서 좀더 거기에 이렇게 depend 하는 것 같아요. 적어도 corpus를 찾아보면 확인해 봤으니까 맞다는 생각이 들어요. 그런 것 때문에 그런 것 같아요. 절적으로 좀 다른…(1/2/04)

My approach to writing did not change all together from a big picture, but since the part of checking the corpus came in my writing process, the quality and process of my writing has changed to a certain degree. In terms of the overall process, my writing process is the same: I finish the whole writing, print it out and revise it. However, I came to check the corpus quite often during writing. I check the corpus to find the most appropriate expressions and depend on the corpus before I produce the final product. If I check the corpus, I feel more confident about the quality of my writing.

3) Collocation 대한 감이 늘었다고 할 수 있는 것 같아요. 근데 그런 이걸 봤기 때문에 그런 건 아니구요, 그냥 이걸 봐서 전반적인 collocation 이나 형태들에 대한 감이 늘어난 것 같구요.
I think my collocational knowledge has been increased. But I don’t think it is only or directly due to corpus use. Rather, after using it, I started to realize that words have specific collocation patterns and began to notice those patterns. That is, increased collocational knowledge may not have been the first effect of corpus use, but a secondary effect.

3. June

1) 거기에서 패턴을 찾는 거. 빈뜻수라든지, 그런 걸 내가 따로 prototype 을 만들어야 되잖아요. 익숙하기 전까지는 그것 자체가 시간이 걸리니까 누군가가 그 단어처럼 “make use of” 단어처럼 해줬으면 좋죠. 처음에 제가 그거를 찾아야되니까, 그 시간만 단축되면. 효율성은 누군가 제시해 주죠. 이렇게 많이 쓰다 숙어로서, 그게 익어요. 처음.. 제가 많이 쓰고 훈련되어 있는 상태라면만 분석이 가능한데 지금 상태에서는 안 되요. 누군가 없어서 이건 이런 거 이렇게 쓰다고 해 주는 게 가장 빠름데, 언제나 그런 분이 있는 게 아닙니까. 혼자서 하는 게 아직까지는 제가 그 정도 수준은 아니라고 생각하니까 그런 것 같아요. (12/22/03)

Finding a pattern and making a prototype from the corpus results takes me too much time. I hope somebody can do it for me and just tell me the common usage patterns, like “make use of.” If I am fully trained in using the corpus, I may be able to analyze the data, but I can’t do it. I think it would be the most time-efficient if somebody can give me the analysis and common patterns. I don’t think I can do it by myself.

2) 특히 단어를 모르면, 제가 말을 하는데 이 말이 생각이 안 나서 말을 못하면 의사 전달이 안 되는 거잖아요. 단어를 어느 정도 알면 문법이 들어와도 이제 그 배열이 되니까, 서로 상관관계는 있거든요. 근데 아까 말한 collocation 이란 개념과 연관을 시켜볼 때 따로 따로 된 단어들이 어떤 문법적인 구조를 가져서, 근데 그게 문법은 사회규약이나 관습에 영향을 받는 거라 생각을 하기 때문에, 근데 그게 먼저 언어학자들이 배열한 거야 되니까 그래서 사람들이 따라서 쓰는 걱정 쓰는 거는 아니라고 생각하세요. 오랫동안 변했을 거고, 그 패턴을 보고 이제 언어학자들이 문법은 이러하다고 해서 이런 이런 단어들끼리 조합을 해서 써야지 서로 다른 사람들끼리 이해할 수 있다 이런 식으로 해서 서로 단어와 단어 짝이 있는 거 같아요. (3/10/04)

If we don’t know words, we can’t communicate at all. If we know words, then grammar comes in, and the words are arranged in order and become interrelated. Collocation seems to be the combination of words and grammar. Words have grammar that is influenced by social customs or regulations. I don’t believe that grammarians first generate rules and we follow them. Rather, language has been evolved and changed
during a long period of time, and grammarians describe the patterns to show us that some words should be used with certain words in order to make others understood.

4. Sung

1) 주로 막히는 부분들이 의미를 모르는 부분도 있고, 보다는 이제 용례를 잘 몰라서, 용례를 알고 싶을 때. 고런 부분은 처음부터 corpus 를 생각하죠. 고럴 때는 사전보다는 corpus 를 먼저 생각하죠. (2/5/04)

Now problems are more about usage rather than meanings, so I think about using the corpus first rather than dictionary. I am thinking about the corpus first for word usage.

2) Writing은 훨씬 편해졌어요. Writing 능력이 갑자기 늘어난 것은보다는 paraphrase를 많이 하게 됐거든요. 그런 면에서 많이 편해졌고 자료를 좀 필요한 부분은 text를 모아줬다가 카드를 가지고 쓰다 보니가 skill이 좀 늘는 것 같어요. 시간도 빨라지고 정확해지고 그리고 거기에 맞는 term을 많이 쓰고. 옛날에는 그런 것 없이 무조건 생각나는 대로 쓰니까 문법적인 면도 틀리고 내용도 이해가 안되고 용어도 많이 부정확했는데. 그런 면에서 이제 많이 좋아졌죠. (3/9/04)

I felt much more comfortable in writing. It’s not because my writing ability increased tremendously all of sudden, but because I learned how to paraphrase texts. I started copying texts from my reading for my own purpose and refer to them. By doing so, my skill in using sources increased, composing gets faster, and writing becomes more accurate by using accurate grammar and terminology.
## APPENDIX I

### SURVEY RESULTS

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean*</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning the usage of phrases</td>
<td>5.23</td>
<td>.73</td>
</tr>
<tr>
<td>Improving English writing skill</td>
<td>5.08</td>
<td>.64</td>
</tr>
<tr>
<td>Improving English academic writing</td>
<td>5.08</td>
<td>.64</td>
</tr>
<tr>
<td>Learning the usage of vocabulary</td>
<td>5.00</td>
<td>.91</td>
</tr>
<tr>
<td>Increasing confidence about English writing</td>
<td>4.92</td>
<td>.76</td>
</tr>
<tr>
<td>Helpful than a dictionary for English writing</td>
<td>4.23</td>
<td>.93</td>
</tr>
<tr>
<td>Learning grammar</td>
<td>4.00</td>
<td>1.08</td>
</tr>
<tr>
<td>Learning the meaning of vocabulary</td>
<td>3.23</td>
<td>.73</td>
</tr>
<tr>
<td>Improving English reading skill</td>
<td>3.08</td>
<td>1.04</td>
</tr>
</tbody>
</table>

* 1=very unhelpful, 2=unhelpful, 3=somewhat unhelpful, 4=somewhat helpful 5=helpful, 6=very helpful

Table I.1: The class’ corpus use, advantages

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean*</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-consuming data analysis</td>
<td>3.69</td>
<td>1.32</td>
</tr>
<tr>
<td>Concordance output analysis problems</td>
<td>3.15</td>
<td>1.14</td>
</tr>
<tr>
<td>Cut-off sentences in concordance output</td>
<td>3.08</td>
<td>1.26</td>
</tr>
<tr>
<td>Collocate output analysis problems</td>
<td>3.08</td>
<td>1.12</td>
</tr>
<tr>
<td>Search technique problems</td>
<td>2.92</td>
<td>1.26</td>
</tr>
<tr>
<td>Internet connection too slow</td>
<td>2.92</td>
<td>1.32</td>
</tr>
<tr>
<td>Computer / Internet accessibility</td>
<td>2.92</td>
<td>1.71</td>
</tr>
<tr>
<td>Too few sentences in concordance output</td>
<td>2.77</td>
<td>1.01</td>
</tr>
<tr>
<td>Unfamiliar vocabulary on output text</td>
<td>2.62</td>
<td>.65</td>
</tr>
<tr>
<td>Real texts too difficult</td>
<td>2.31</td>
<td>.75</td>
</tr>
<tr>
<td>Too many sentences in concordance output</td>
<td>2.23</td>
<td>.93</td>
</tr>
<tr>
<td>Content because it doesn’t apply to student’s field</td>
<td>2.00</td>
<td>.58</td>
</tr>
</tbody>
</table>

* 1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree 6=strongly agree

Table I.2: The class’ corpus use, problems/difficulties
<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prototype construction is a useful language learning experience</td>
<td>4.92</td>
<td>.76</td>
</tr>
<tr>
<td>I search the corpus for writing problems</td>
<td>4.38</td>
<td>1.33</td>
</tr>
<tr>
<td>I usually get the information I need from the search</td>
<td>4.15</td>
<td>.69</td>
</tr>
<tr>
<td>Was helpful for revising papers for this course</td>
<td>4.15</td>
<td>.90</td>
</tr>
<tr>
<td>Was helpful for composing papers for this course</td>
<td>4.15</td>
<td>1.14</td>
</tr>
<tr>
<td>Prototype construction is easy</td>
<td>4.15</td>
<td>.99</td>
</tr>
<tr>
<td>I often use the corpus for my L2 writing</td>
<td>4.08</td>
<td>1.12</td>
</tr>
<tr>
<td>I feel confident in prototype construction</td>
<td>4.00</td>
<td>.82</td>
</tr>
<tr>
<td>I often use the corpus of my own choice</td>
<td>3.92</td>
<td>1.19</td>
</tr>
<tr>
<td>I use the corpus for other courses too</td>
<td>3.85</td>
<td>1.07</td>
</tr>
<tr>
<td>I use the class strings as a reference in writing</td>
<td>3.62</td>
<td>1.26</td>
</tr>
</tbody>
</table>

*1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree 6=strongly agree

Table I.3: The class’ responses to corpus use in ESL writing instruction

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would recommend corpus use in the same future course</td>
<td>5.08</td>
<td>.49</td>
</tr>
<tr>
<td>Would recommend corpus use to other international students</td>
<td>5.00</td>
<td>.41</td>
</tr>
<tr>
<td>The corpus is a useful resource for L2 writing</td>
<td>5.00</td>
<td>.58</td>
</tr>
<tr>
<td>Will use the corpus for future writing</td>
<td>4.92</td>
<td>.49</td>
</tr>
<tr>
<td>Would recommend corpus use to home country</td>
<td>4.69</td>
<td>.75</td>
</tr>
<tr>
<td>I like it more as I use it more</td>
<td>4.15</td>
<td>.80</td>
</tr>
<tr>
<td>I feel comfortable using corpus technology</td>
<td>4.00</td>
<td>.82</td>
</tr>
<tr>
<td>Corpus use changed my approach to L2 writing</td>
<td>4.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Corpus use changed my understanding of the language</td>
<td>3.77</td>
<td>1.01</td>
</tr>
<tr>
<td>It would have been helpful for a TOEFL test</td>
<td>3.77</td>
<td>1.17</td>
</tr>
<tr>
<td>Corpus use changed my L2 writing process</td>
<td>3.69</td>
<td>.95</td>
</tr>
</tbody>
</table>

*1=strongly disagree, 2=disagree, 3=somewhat disagree, 4=somewhat agree, 5=agree 6=strongly agree

Table I.4: The class’ overall evaluations of corpus use in L2 writing