WEB-BASED CONCORDANCING AND OTHER REFERENCE RESOURCES AS A PROBLEM SOLVING TOOL FOR L2 WRITERS: A MIXED METHODS STUDY OF KOREAN ESL GRADUATE STUDENTS’ REFERENCE RESOURCE CONSULTATION

by

Choongil Yoon

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Choongil Yoon 2014
Abstract

The present study investigated how 6 Korean graduate students at a Canadian university used a suite of multiple Web-based reference resources (named i-Conc), consisting of concordancers and dictionaries, as a cognitive tool for solving linguistic problems encountered over the course of completing—in English, their second language (L2)—an academic writing assignment for one of their graduate courses. Using a mixed methods design employing surveys, interviews, screen recordings, a query tracking log, and detailed case studies, the thesis provides rich descriptions of (a) the processes, and outcomes of the 6 participants’ uses of i-Conc as a reference tool for their writing authentic academic tasks and (b) their perceptions of the suite as a means of writing assistance.

Overall, i-Conc served as an intellectual partner that aided the participants in strategically solving lexical and grammatical problems during their writing assignments: About 70% of the problems they addressed with i-Conc resulted in correct text formulations or revisions. The different resources in i-Conc were each shown to have unique functions for which they were best suited, suggesting that concordancing may optimally be consulted in combination with, not in place of, other resources. The benefits of consulting i-Conc for L2 writing went beyond simply helping the participants’ problem solving to potentially facilitating
their language acquisition. Input-feedback interactions with the reference suite prompted the
participants to carry out robust meaning negotiations in their efforts to verify their intuitive
hypotheses and to venture beyond their current linguistic repertoires.

Participants acted on these potential benefits somewhat differently. Case studies and
cross-case analyses demonstrated complex interactions between the participants’ individual
traits and goals, the educational contexts for which they were writing, and their perceptions and
evaluations of particular affordances provided by i-Conc. These findings imply that to build
meaningful cognitive partnerships with reference tools, L2 writers should receive progressive
guidance on principles for effective reference resource consultation along with training in
strategies for using different types of resources, contingent on individuals’ abilities and ongoing
needs arising from their macro and micro contexts for writing and for language learning.
Acknowledgements

My name is printed in bold on the cover, but this doctoral thesis is not solely mine. There are many people who were instrumental and indispensable in bringing the thesis to fruition.

First, I would like to express the deepest appreciation to Dr. Alister Cumming, my academic advisor and thesis supervisor. His guidance and mentorship was a steady source of inspiration for the development of the theoretical and methodological basis of the study and the organization and editing of the thesis. I am also deeply grateful to Dr. Eunice Jang, and Dr. Jim Cummins, two other members of my dissertation committee, who provided me with useful suggestions and insightful feedback to drafts of my proposal and thesis. I would also like to thank Dr. Carol Chapelle and Dr. Marlene Scardamalia for their thoughtful feedback and suggestions for future direction at my final oral examination.

I would like to express my deepest gratitude as well to my participants for patiently carrying out all the tasks despite their busy schedules and sharing their experiences and insights with sincerity. While I tried to help their English academic writing with i-Conc, I learned as much from them and had a precious opportunity to grow as a teacher and as a researcher.

I also owe a great debt of gratitude to Hwangjoon Kwon, my dear friend and former colleague, who developed i-Conc program for the study and provided technical advice and support throughout data collection and analysis. His help was one major contributor to making this study and thesis possible.

My special thanks goes to Hyeeyoon Cho, and Paul Quinn, who kindly helped me code the data and made valuable suggestions. I am also thankful to Myeonghyeon, Seungwon, Hyuna, Tim, Daphne, and other friends and colleagues for their emotional support and intellectual stimulation throughout my graduate years.

Lastly, but most importantly, I would like to send Heesoon, my mom, my deepest love and appreciation for being a constant source of encouragement and willpower to forge ahead in the moments of frustration, disappointment, and setbacks during my studies. I dedicate this thesis to my mom.
## Table of Contents

ABSTRACT………………………………………………………………………..…………………………………ii
ACKNOWLEDGEMENTS……………………………………………………………………………………..iv
LIST OF TABLES…………………………………………………………………………………………………………………ix
LIST OF FIGURES……………………………………………………………………………………………………………….....x
LIST OF APPENDICES……………………………………………………………………………………………………..xi

Chapter
Page

1 INTRODUCTION ....................................................................................................................... 1
  1.1 Learner concordancing and language learning............................................................... 2
  1.2 Concordancing and L2 writing ....................................................................................... 4
  1.3 Reference tools in L2 writing ....................................................................................... 6
  1.4 Statement of the problem ......................................................................................... 7
  1.5 The present study ..................................................................................................... 9
      1.5.1 Purpose and design ............................................................................................ 9
      1.5.2 Research questions ......................................................................................... 10
      1.5.3 Significance of the study .............................................................................. 11

2 REVIEW OF RELEVANT THEORIES AND RESEARCH ..................................................... 13
  2.1 Corpus linguistics and its contributions to L2 pedagogy ........................................... 13
  2.2 Theories underpinning learner concordancing for L2 writing ................................ 15
      2.2.1 Constructivist learning theory ........................................................................ 15
      2.2.2 Data-driven learning (DDL) ......................................................................... 16
      2.2.3 Cognitive tools and distributed cognition ..................................................... 18
      2.2.4 Affordance .................................................................................................. 20
  2.3 Empirical studies of concordancing as a reference resource ..................................... 21
      2.3.1 Concordancing for error correction and revision ........................................... 21
      2.3.2 Longitudinal concordancing studies .............................................................. 25
      2.3.3 Google as writing assistance ...................................................................... 27
  2.4 Concordancing as a problem solving tool ................................................................. 30
      2.4.1 Types of problems arising during text formulation ....................................... 30
      2.4.2 Types of problems addressed with corpus tools ......................................... 32

3 METHODS ................................................................................................................................ 34
  3.1 Participants .............................................................................................................. 34
  3.2 Reference suite (i-Conc) .......................................................................................... 36
      3.2.1 Corpus of Contemporary American English (COCA) ................................ 38
      3.2.2 Google ........................................................................................................ 40
      3.2.3 Just The Word ............................................................................................ 43
      3.2.4 Dictionaries ............................................................................................ 45
  3.3 Procedures ............................................................................................................... 45
      3.3.1 Training ................................................................................................... 46
3.3.2 Data collection

3.3.2.1 Surveys

3.3.2.2 Interviews

3.3.2.3 Writing assignments

3.3.2.4 Query log

3.3.2.5 Screen recording and stimulated recall

3.4 Data analysis

3.4.2 Coding scheme – types of problems and queries

3.4.3 Coding scheme – problem solving results

3.4.4 Case studies and cross-case analysis

3.5 Reliability/credibility

3.6 Notation convention

4 GROUP FINDINGS

4.1 Profiles of the participants

4.1.1 Experience and perceptions of English academic writing

4.1.2 Writing support and reference resource use

4.2 How and for what participants consulted i-Conc

4.2.1 Typical sequence of i-Conc consultation

4.2.2 Types of problems and queries

4.2.2.1 Analysis of problems

4.2.2.2 Analysis of queries

4.2.3 Resources consulted

4.3 Effects of using i-Conc on problem solving performance

4.3.1 Brief overview of different code sets

4.3.2 Findings from each combination of groups

4.4 Strategies and pitfalls in i-Conc consultation

4.4.1 Strategy use in i-Conc consultation

4.4.1.1 Query strategies

4.4.1.2 Evaluation/application strategies

4.4.2 Pitfalls in i-Conc consultation

4.4.2.1 Lack of rigor in evaluating query results

4.4.2.2 Lack of flexibility in trying out different options

4.4.2.3 Limitations of resources

4.5 Participants’ evaluations of i-Conc

4.5.1 Benefits and challenges of using i-Conc

4.5.2 Changes in confidence, attitudes, and approaches

5 CASE STUDIES

5.1 Case study 1: Jae

5.1.1 Background

5.1.2 i-Conc use

5.1.3 Evaluation of i-Conc

5.2 Case study 2: Yumee

5.2.1 Background

5.2.2 i-Conc use
5.2.3 Evaluation of i-Conc ................................................................. 119
5.3 Case Study 3: Jinho ................................................................. 120
  5.3.1 Background ........................................................................ 120
  5.3.2 i-Conc use ........................................................................ 121
  5.3.3 Evaluation of i-Conc .......................................................... 125
5.4 Case Study 4: Goeun ............................................................ 126
  5.4.1 Background ........................................................................ 126
  5.4.2 i-Conc use ........................................................................ 127
  5.4.3 Evaluation of i-Conc .......................................................... 129
5.5 Case Study 5: Shia ................................................................. 131
  5.5.1 Background ........................................................................ 131
  5.5.2 i-Conc use ........................................................................ 132
  5.5.3 Evaluation of i-Conc .......................................................... 134
5.6 Case Study 6: Ian ................................................................. 136
  5.6.1 Background ........................................................................ 136
  5.6.2 i-Conc use ........................................................................ 137
  5.6.3 Evaluation of i-Conc .......................................................... 139

6 CROSS-CASE ANALYSIS: INTERVENING FACTORS INFLUENCING I-CONC USE . . 142
  6.1 Writer factors ........................................................................ 143
  6.2 Text factors ............................................................................ 147
  6.3 Context factors ..................................................................... 149
  6.4 Summary across cases .......................................................... 150

7 DISCUSSION AND IMPLICATIONS .............................................. 152
  7.1 Major findings ...................................................................... 152
    7.1.1 Problems and queries ..................................................... 152
    7.1.2 Resources consulted ..................................................... 154
    7.1.3 Problem-solving performance ...................................... 158
    7.1.4 Strategies and pitfalls ................................................... 159
    7.1.5 Participants’ evaluations of i-Conc .................................. 162
    7.1.6 Individual differences and intervening factors ............. 165
  7.2 Implications .......................................................................... 166
    7.2.1 Implications for theories: i-Conc as a cognitive tool .... 166
      7.2.1.1 Cognitive effects of i-Conc ..................................... 167
      7.2.1.2 Cognitive extension vs. cognitive overload/distraction . 168
      7.2.1.3 Potential role of i-Conc for language acquisition through L2 writing . 170
    7.2.2 Pedagogical implications ............................................... 172
      7.2.2.1 Implications for learners ....................................... 172
      7.2.2.2 Implications for teachers ...................................... 173
      7.2.2.3 Implications for learner training .......................... 174
      7.2.2.4 Implications for reference tools ........................... 178
  7.3 Limitations of the study and directions for future research .. 179
    7.3.1 Limitations of the study ............................................... 180
    7.3.2 Suggestions for future studies ..................................... 181
List of Tables

Table 3.1  Participant Profiles ..........................................................................................35
Table 3.2  Reference Resources Featured in \( i\)-Conc ..........................................................38
Table 3.3  Custom Search Engines Featured in \( i\)-Conc ......................................................42
Table 3.4  Procedures ........................................................................................................46
Table 3.5  Content of Tutoring ..........................................................................................48
Table 3.6  Writing Assignments .........................................................................................50
Table 3.7  Length of Participants’ Screen Recording .........................................................53
Table 3.8  Coding Scheme for Problems and Queries .......................................................56
Table 3.9  Coding Scheme for Problem Solving Results ..................................................59
Table 4.1  Participants’ Prior Experience of Academic Writing ........................................63
Table 4.2  Sources of Writing Support and Reference Resources ....................................67
Table 4.3  Number of Queries Performed by the Participants ............................................74
Table 4.4  Major Query Purposes .......................................................................................79
Table 4.5  Types and Linguistic Categories of the Participants’ Queries ...............................81
Table 4.6  The Number and Percentage of Problem Spaces Belonging to each Group ....88
Table 4.7  Classification of Problem Spaces According to their Problem Solving Results .................................................................................................................................89
Table 4.8  Effects of \( i\)-Conc Consultation by Problem Type and Primary Query Purpose .................................................................................................................................90
Table 4.9  Main Query Strategies and Examples ................................................................94
Table 4.10 Evaluation/application Strategies and Examples .............................................96
Table 4.11 Participants’ Perceived Usefulness and Difficulty of Use of \( i\)-Conc Resources .................................................................................................................................104
Table 4.12 Summary of Participants’ Evaluations of \( i\)-Conc .............................................107
Table 5.1  An overview of Jae’s \( i\)-Conc use .....................................................................109
Table 5.2  An overview of Yumee’s \( i\)-Conc use .................................................................116
Table 5.3  An overview of Jinho’s \( i\)-Conc use .................................................................117
Table 5.4  An overview of Goeun’s \( i\)-Conc use ...............................................................127
Table 5.5  An overview of Shia’s \( i\)-Conc use .................................................................132
Table 5.6  An overview of Ian’s \( i\)-Conc use .................................................................137
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>KWIC concordance output of the word “example”</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Types of problems that arise during text formulation</td>
<td>31</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Functions of reference resource consultation</td>
<td>33</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Screenshot of i-Conc</td>
<td>37</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Screenshot of LIST display on COCA</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Screenshot of KWIC display on COCA</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Screenshot of the Google tab on i-Conc</td>
<td>41</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Screenshot of the results of the query “assumption” run in JTW combinations</td>
<td>43</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Screenshot of the results of the query “make assumption” run in JTW alternatives</td>
<td>44</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Screenshot of concordances of the word combination “make assumption”</td>
<td>45</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Screenshot of the Dictionaries tab on i-Conc</td>
<td>45</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Screenshot of query log</td>
<td>51</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>Screenshot of a participant’s screen-recorded video</td>
<td>53</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Reference resource consultation models in previous studies</td>
<td>69</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>i-Conc consultation sequence</td>
<td>70</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>An example of i-Conc consultation sequence</td>
<td>72</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Frequencies of initial query purposes</td>
<td>76</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Frequencies of query purposes</td>
<td>77</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Examples of each query type</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Frequencies of query purposes, types and linguistic categories in each resource</td>
<td>83</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Frequencies of resource use for major query purposes</td>
<td>85</td>
</tr>
<tr>
<td>Figure 6.1</td>
<td>Distributions of problem and query types in each participant’s i-Conc use</td>
<td>143</td>
</tr>
<tr>
<td>Figure 6.2</td>
<td>Percentages of SC queries and Google use in each participant’s i-Conc use</td>
<td>143</td>
</tr>
<tr>
<td>Figure 6.3</td>
<td>Percentages of satisfactory consultations and successful resolutions in each participant’s total problem spaces</td>
<td>145</td>
</tr>
<tr>
<td>Figure 6.4</td>
<td>Percentage of Equi queries in each participant’s total queries</td>
<td>148</td>
</tr>
<tr>
<td>Figure 6.5</td>
<td>Factors influencing the participants’ i-Conc use and its evaluation</td>
<td>150</td>
</tr>
</tbody>
</table>
# List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>Information Letter/Consent Form</td>
<td>196</td>
</tr>
<tr>
<td>Appendix B</td>
<td>i-Conc Tutorial</td>
<td>199</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Questionnaires</td>
<td>211</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Interview Guide</td>
<td>215</td>
</tr>
<tr>
<td>Appendix E</td>
<td>Stimulated Recall Guide</td>
<td>218</td>
</tr>
<tr>
<td>Appendix F</td>
<td>An Example of a Problem pace</td>
<td>219</td>
</tr>
<tr>
<td>Appendix G</td>
<td>Possible Code Sets and the Solution Path Each Code Set Represents</td>
<td>220</td>
</tr>
<tr>
<td>Appendix H</td>
<td>Principles of Effective Reference Resource Consultation</td>
<td>221</td>
</tr>
</tbody>
</table>
CHAPTER ONE

INTRODUCTION

Over the past few decades, the rapid spread of personal computers and innovations in Internet technology have been changing the processes by which language learners go about writing in a foreign or second language (L2) (Stapleton & Radia, 2009; Warschauer, 2007). One of the changes this digital age has brought about is a proliferation of online reference resources that provide support for L2 writers. Specifically, advances in data processing power and storage capacity have not only made traditional reference resources such as dictionaries, and thesauri available online—with ample and various linguistic information that cannot be provided in offline resources for lack of space—but also generated new types of reference resources with a great potential for supporting L2 writers (Frankenberg-Garcia, 2005; Tono, 2012).

Learner concordancing, or learners’ direct use of corpora—one of these new reference resources—has particularly been suggested as a useful tool that provides language learners with typical and frequent patterns in which a linguistic item is used (Bloch, 2007; Chambers, 2005; Hyland, 2003; Johns, 1988, 1991; Stapleton & Radia, 2009). Since the 1990s, concordancing has been steadily explored as a language pedagogy tool, drawing in large part on what Johns (1988, 1991) introduced as data-driven learning (DDL). However, it is only relatively recently that corpus consultation has started to be empirically examined as a reference tool for L2 writing (Chambers & O’Sullivan, 2004; Gaskell & Cobb, 2004; Kennedy & Miceli, 2001, 2010; O’Sullivan & Chambers, 2006; Park, 2010; Sun, 2007; H. Yoon, 2005, 2008; Yoon & Hirvela, 2004). These studies, on the whole, conclude that with the right amount of training and motivation, concordancing can be an effective tool to support learners in solving lexical and grammatical problems encountered while writing in their target language.

However, the previous studies on concordancing as a reference tool have been, in large part, confined to examining one-off corpus tool uses for written error correction or translation tasks, mostly conducted in classroom settings, and often with corpus tools provided as the only reference resources to consult. These constrained task settings, though necessary for keeping variance to a minimum for research purposes, may not reflect typical and authentic writing tasks in a university setting (Murphy & Roca de Larios, 2010; Stapleton, 2010).
The aim of the present study is to examine how and to what extent L2 writers would consult newly emerging reference resources in combination with more traditional online resources in a naturalistic university setting. Specifically, the study examined L2 writers’ reference-resource consultation while independently engaging in an authentic writing assignment (e.g., a course or research paper) completed over several weeks. A particular focus was placed on the role of concordancing and its combined use with other resources as a cognitive tool that amplifies and extends the L2 writer’s cognitive powers in solving the immediate linguistic problems that arise while writing an academic paper in an L2. In what follows, the background of and justification for the study is discussed.

1.1 Learner concordancing and language learning

Over the past few decades, insights from corpus linguistics have been applied to language teaching to a large extent in indirect manners informing the design and development of dictionaries, grammar books, syllabuses and teaching materials, which are characterized by emphases on authentic language examples, frequency and typicality, and the close relationship between lexis and grammar. More recently, corpora and concordancing tools have started to be directly used by language teachers and learners themselves.

In corpus linguistics, a corpus is usually defined as a collection of naturally occurring texts compiled for linguistic study (Biber, Conrad, & Reppen, 1998), and concordancing refers to searching a corpus and producing a list of instances of a given linguistic item (a word or phrase) using computer software called a concordancer. In its most common format or keyword in context (KWIC), a concordancer displays a given linguistic item in the center together with its immediate context on both sides or co-text (see Figure 1.1 for an example of KWIC concordance output). Concordance outputs provide, horizontally read, information on how particular items are used in context (i.e., language use) and, vertically read, repeated patterns associated with those items (i.e., evidence of system) (Hyland, 2003; Tognini Bonelli, 2010)
Many researchers have agreed that there are a number of benefits in using concordancing for language teaching and learning in general (Bloch, 2007; Conrad, 2000, 2008; Chambers, 2005; Granath, 2009; Hyland, 2003; Johns, 1988; O’Sullivan, 2007; Stapleton & Radia, 2009). Firstly, by providing frequency information, corpora provide learners with the most common and typical language forms that occur in a given context. Secondly, learners can have a rich experience of “real language,” or language used in authentic contexts as opposed to invented or contrived examples in non-corpus based grammar books. Thirdly, concordance examples show the functions of words in different contexts and genres, and thereby raise learners’ language awareness. Fourthly, learners can access the lexicogrammatical patterns and phraseologies (collocation, colligation, etc.) that cannot be made visible (or not as effectively) through other resources such as dictionaries or a native speaker’s intuition. Finally, drawing on all the aforementioned benefits, corpus consultation by learners can foster inductive learning through analyzing a large amount of language examples. Inductive learning through corpus analysis in particular is often linked to what Johns (1988, 1991) introduced as data-driven learning (DDL). In DDL, the teacher as a coordinator engages learners as researchers in the analysis of corpus data and encourages them to make their own discoveries about how language works in ways that may not be possible without the use of corpora. In other words, learners are encouraged to learn from examples provided by a concordancer rather than categories and definitions provided in dictionaries or grammar books.

Figure 1.1. KWIC concordance output of the word “example”.

Many researchers have agreed that there are a number of benefits in using concordancing for language teaching and learning in general (Bloch, 2007; Conrad, 2000, 2008; Chambers, 2005; Granath, 2009; Hyland, 2003; Johns, 1988; O’Sullivan, 2007; Stapleton & Radia, 2009). Firstly, by providing frequency information, corpora provide learners with the most common and typical language forms that occur in a given context. Secondly, learners can have a rich experience of “real language,” or language used in authentic contexts as opposed to invented or contrived examples in non-corpus based grammar books. Thirdly, concordance examples show the functions of words in different contexts and genres, and thereby raise learners’ language awareness. Fourthly, learners can access the lexicogrammatical patterns and phraseologies (collocation, colligation, etc.) that cannot be made visible (or not as effectively) through other resources such as dictionaries or a native speaker’s intuition. Finally, drawing on all the aforementioned benefits, corpus consultation by learners can foster inductive learning through analyzing a large amount of language examples. Inductive learning through corpus analysis in particular is often linked to what Johns (1988, 1991) introduced as data-driven learning (DDL). In DDL, the teacher as a coordinator engages learners as researchers in the analysis of corpus data and encourages them to make their own discoveries about how language works in ways that may not be possible without the use of corpora. In other words, learners are encouraged to learn from examples provided by a concordancer rather than categories and definitions provided in dictionaries or grammar books.
There are a wide variety of freely and commercially available corpora and concordancing tools (or concordancers) that come in different types and with different functionalities (see Römer, 2006, for an overview of major corpus/concordancer resources). However, the corpora used in research and classroom applications are by and large of two types (see Boulton, 2010b for a comprehensive review of concordancing studies). First, many studies have used large general corpora—which are balanced in terms of domains the language data are compiled from, and tagged for part-of-speech (POS) information—such as the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA), or subsections of them. Second, in other studies, specialized corpora were compiled by the researchers or teachers from only the text types and genres relevant to the target learners. Concordancers used in previous studies and applications can also be broadly broken down to standalone concordancing programs and Web-based concordancing interfaces. Most studies have commonly showed that as a newly emerging resource involving computer use and specific skills, concordancing requires long preparation and gradual, guided training on the part of teachers and learners alike.

More recently, in what is broadly defined as Google-assisted language learning (GALL), Internet search engines have also been suggested and explored as concordancing tools (Conroy, 2010; Fujii, 2007; Krajka, 2009; Kuo, 2008; Sha, 2010; Shei, 2008). Most GALL approaches center around the use of the Web as a corpus and Google (or other search engines) as a matching concordancer. These studies on the whole have showed that Web concordancing has advantages over corpus resources in the traditional sense mentioned above. Chief among them is learners’ familiarity with search engines. Today’s language learners are “digital natives” (Palfrey & Gasser, 2008), who have long been using search engines (represented by Google) not only to get content information but also to test their hypotheses about L2 forms (Conroy, 2010; Hafner & Candlin, 2007). Therefore, they are already familiar with the basic functions of search engines and have developed their own strategies for their effective use.

1.2 Concordancing and L2 writing

In reference to L2 writing, Hyland (2003) classified the uses of concordancing into two categories depending on the role played. Concordancing as a research tool is intended to help L2 learners to discover underlying rules and regularities about a given linguistic item from its
concordance. Concordancing in this role is how concordancing has typically been defined and practiced in the majority of previous research and classroom applications (Boulton, 2011). When it comes to L2 writing studies, concordancing has been explored mostly in the form of DDL activity in a language or linguistics course, often with explicit language learning goals (e.g., learning specific lexical, grammatical or genre items, or concordancing itself) (e.g., Cresswell, 2007; Lee & Swales, 2006; Turnbull & Burston, 1998).

Concordancing can also be used as a reference tool that L2 writers consult to fill knowledge gaps they encounter while writing in the target language. The use of concordancing here usually does not involve learning specific linguistic items or language learning per se. Rather, it serves as more of a problem solving tool that aids L2 learners in successful completion of their writing tasks. Many of recent studies on the effects of concordancing on L2 writing examined this line of concordancing use (Frankenberg-Garcia, 2005; Kennedy & Miceli, 2001, 2010; Park, 2010; H. Yoon, 2005, 2008).

However, these two uses do not exclude each other. Given the relationship between learning and problem solving, any form of concordancing use would involve varying degrees of both types. As a matter of fact, most studies of corpus tool use for learners’ written error corrections or revisions involved roughly equal degrees of both uses as the tasks conducted were designed with explicit language learning goals (Chambers & O’Sullivan, 2004; Gaskell & Cobb, 2004; Gilmore, 2009; O’Sullivan & Chambers, 2006; Watson Todd, 2009).

Although often mixed, the results of these studies (see C. Yoon, 2011 for an overview of studies on concordancing in L2 writing) indicate that while learner concordancing can be an effective tool that provides L2 writers with evidence of language use in the forms of lexi-grammatical patterns and frequency information, the success and motivation of L2 writers in its use can vary depending on a wide variety of factors such as language proficiency, learning style, task, and the actual tools used. Nonetheless, a synthesis of previous research points to an overall trend that, as with other pedagogical tools, L2 writers tend to be more motivated to use concordancing for tasks that are directly linked to their needs such as finding solutions to authentic problems at hand (Lee & Swales, 2006; Watson Todd, 2001), and they actually use it more effectively in such situations than when they simply carry out concordancing tasks for purposes imposed by teachers. That is, the potential of concordancing for supporting L2 writers can be more directly and immediately realized when it is used as a reference tool that learners
consult to address the accuracy and appropriacy of the lexical and grammatical elements of their writing.

1.3 Reference tools in L2 writing

Accuracy and appropriacy of the lexical and grammatical features of writing become increasingly important as L2 writers advance to higher levels in their studies because greater articulation is needed for accurately conveying increasingly complex content and ideas. However, previous research has shown that while well-trained in idea development and organization through years of instructions and experience, L2 learners even at advanced levels still struggle with the linguistic aspects of their writing such as lexis and grammar (Hinkel, 2002; Silva, 1993; H. Yoon, 2005, 2008). Specifically compared to writing in their L1, learners devote much more time and attention to lexical searches and syntactic encoding that will accurately express their ideas into L2 text (Cumming, 2001; Murphy & Roca de Larios, 2010).

Although accuracy/appropriacy in written form is still given great importance in the professional and academic world and is often a major factor that affects instructors’ assessments of students’ writing (Hegelheimer, 2006; H. Yoon, 2008), writing support for L2 writers in these aspects is far from sufficient. Instructors’ feedback on lexical and grammatical features in L2 writers’ papers can hardly be expected in content courses at universities. Even in academic writing courses and writing centres, lexical and grammatical accuracy tends to be neglected, particularly at graduate levels, in favor of idea development and text organization under the influences of process-oriented writing pedagogy (Canagarajah, 2002; H. Yoon, 2005). Furthermore, even the support L2 writers can receive from these sources, if any, is provided mostly after or near the completion of a writing assignment but not during composition.

In this regard, language reference resources (e.g., dictionaries) that L2 writers consult while composing and revising their writing are often the only sources from which they can get immediate support for solving linguistic problems as they arise. Indeed, with the rapid advances in corpus-informed lexicography and text processing technologies over the past decades, a wide variety of reference resources have become available, which include not only corpus-based learners’ dictionaries, collocation dictionaries, and now concordancing tools but also new media
such as CD ROMs, stand-alone software, and online resources (Frankenberg-Garcia, 2011; Tono, 2010). Today’s learners thus have wider choices of reference resources than ever before.

However, research suggests that this choice is to a large extent greatly underexploited by L2 writers, whose reference resource uses are still confined more or less to L1-L2 equivalent look-ups with bilingual dictionaries (Frankenberg-Garcia, 2005, 2011; Nesi, 2012). Thus, to address the difficulties in terms of accuracy issues discussed above, it is crucial that L2 writers learn different uses and functions of reference resources, choose the reference resources that are best suited for their reference needs, and develop skills and strategies for effective and efficient consultation of these resources (Nesi, 2012). L2 writing research and pedagogy should play a greater role in helping L2 writers do so not only to improve their writing performance but also to increase their confidence and autonomy in L2 writing (East, 2008; Kuo, 2007, H. Yoon, 2008).

1.4 Statement of the problem

Despite its potential for providing L2 writers with immediate support for addressing lexical and grammatical accuracy and appropriacy of their writing, as noted earlier, little research has been conducted that provides a comprehensive picture of how L2 writers independently use concordancing along with other resources for referencing purposes in their authentic academic writing. While a considerable number of the studies have directly and indirectly explored the potential of concordancing as a reference tool aiding L2 writers in solving lexical and grammatical problems arising from their writing, they were understandably limited—as most of them were exploratory in nature—in terms of task settings and data collection methods.

While many of the studies had university students as their participants, the tasks they used were often timed, one-off corpus tool use tasks for written error corrections or revisions, mostly in a classroom setting with explicit language learning goals (Chambers & O’Sullivan, 2004; Gilmore, 2008; Kennedy & Miceli, 2010; O’Sullivan & Chambers, 2006) or for translation (Frankenberg-Garcia, 2005). These types of tasks have, among others, two major limitations. First, they do not reflect the typical and common type of academic writing in which L2 writers engage in a university setting (Murphy & Roca de Larios, 2010; Stapleton, 2010), for
which L2 writers do not necessarily have language learning as their goal. Second, investigating the uses of concordancing only for error corrections or revisions, which are largely done at the later stages of the writing process, may not provide much insight into how L2 writers deal with language problems that arise at the moment of composing. A more naturalistic task setting for university students would then be one in which the L2 writer uses concordancing as a reference tool over the entire course of completing an authentic academic paper (e.g., a term paper for a content course, or a research proposal), for which language learning is not necessarily the main goal.

Certain previous studies also had limitations in terms of the scope and methods of data collection. Most studies that investigated the effects of concordancing on written error corrections and revisions tend to report only the outcomes of participants’ corpus tool use as manifested in their final written products, without providing detailed descriptions of the interactions between the participants and the corpus tools, that is, to what extent, for what purposes, and how participants consulted the corpora while engaging in their writing tasks (Chambers & O’Sullivan, 2004; Gaskell & Cobb, 2004; Gilmore, 2008; O’Sullivan & Chambers, 2006). On the other hand, the studies that did examine and report the processes of research participants’ corpus consultations (Frankenberg-Garcia, 2005; H. Yoon, 2005, 2008) based their analyses, to a large extent, on the participants’ self-reports, and manually-kept search logs. While providing information otherwise unavailable, the data obtained through these methods only provide an incomplete picture of the thought processes the L2 writer actually go through while using a corpus tool. Self-reports or retrospective data can often be selective and incomplete reconstructions of what occurred (Kasper, 1998; Park & Kinginger, 2010), and keeping a corpus search log manually while completing an authentic writing task may cause considerable disruption and therefore often make participants avoid a corpus search or its entry in the log.

Finally, in most of the studies cited above, corpus tools were provided as the only reference resources to consult or encouraged as the main resource to use while performing the given tasks. Given that today’s L2 writers have a wider choice for reference resources than ever before and may already be using some of them, confining participants’ reference resource use to concordancing and especially one specific type of concordancing tool would not reflect this real-life choice L2 writers have (Frankenberg-Garcia, 2005). As recently suggested by some
researchers, it would be more meaningful to examine how concordancing can be used strategically in combination with the existing references that learners have already become familiar with (Conroy, 2010; Flowerdew, 2010; Kaur & Hegelheimer, 2006; Pérez-Paredes, Sánchez-Tornel, & Alcaraz Calero, 2012).

To sum up, most of the previous studies in this area lacked, to varying degrees, ecological validity in their findings due to the limitations of the tasks, tools, or data collection methods used.

1.5 The present study

1.5.1 Purpose and design

To address these gaps identified from previous research, the present study traced how six Korean ESL graduate students in a Canadian university independently used i-Conc, a suite of Internet-based language reference resources (consisting of Web-based concordancers, Google and online dictionaries), for solving linguistic problems over the entire course of completing an authentic writing assignment.

Departing from the general DDL approach adopted in previous learner concordancing studies, which views concordancing as a language learning tool, the present study examined the potential of reference resources as a cognitive tool (Jonassen, 1992; Norman, 1993) that amplifies and extends L2 writers’ cognitive powers and mediates their problem solving while writing in the target language. Specifically, the study set out to describe in detail the extent and processes of the participants’ interactions with the reference tool and ultimately gain insights into how to help L2 writers to effectively use reference resources to meet their specific needs for academic writing and thereby become more autonomous writers.

To investigate these under-researched aspects of learner concordancing in an ecologically valid setting, special care was taken for the research design following some of the principles of design-based research, a methodology that has been increasingly adopted in educational studies that pursue both theory building about and practical applications of given instructional or technological interventions (Anderson & Shattuck, 2012; Wang & Hannafin, 2005). Firstly, the present study situated the use of the reference resources in a naturalistic
context where the participants engaged in one of their “real-world” writing assignments with all their contextual complexities and constraints (as opposed to writing tasks designed for research purposes). Secondly, the study was designed to have two phases. Based on the results of the first exploratory phase, the research instruments, reference suite, and analytic framework were developed, and also refined throughout the second phase on an ongoing basis. Thirdly, the participants were not simply observed for their tool use during their writing tasks, but provided with initial tutoring in the basic concepts around the resources and their effective use at the beginning of the study and with personalized feedback to enhance their tool use throughout the study. Lastly, to capture the processes of the participants’ reference suite consultation in great detail, each participant’s interactions with the suite was traced in a case study using multiple methods such as surveys, screen recordings, a query log, interviews and stimulated recalls, which were in turn intended to enhance the reliability and trustworthiness of the analyses and inferences based on them.

1.5.2 Research questions

The present study has the following specific research questions:

i) How and for what purposes do the participants consult a suite of Internet-based language reference resources (Web-based concordancers and online dictionaries) while they engage in an authentic academic writing assignment? What types of linguistic problems do they address with the reference suite? What reference resources do they consult for different types of problems?

ii) How does the use of the reference suite affect the participants’ problem solving performance?

iii) What strategies do the participants use and develop in their interactions with the reference suite? What difficulties and challenges do they experience?

iv) How do participants evaluate Web-based concordancing and other reference resources in terms of their utility as writing assistance? Particularly, how does the use of the reference suite affect their
approaches, attitudes and confidence with respect to English academic writing?

v) What are the factors that contribute to individual differences, if any, in the participants’ interactions with and evaluations of the suite?

1.5.3 Significance of the study

Despite the recent emphasis on content and idea development in writing under the influences of process-oriented writing pedagogy and research, L2 writers still need support in the language features of their writing in terms of accurate and appropriate use of vocabulary and grammar. While writing support in these areas are not sufficiently available from content courses or the writing centres of North American universities, the addition of concordancing to L2 writers’ existing reference resources can help them to get immediate support for solving language problems that arise while writing in the target language. Most previous research has taken for granted the use of concordancing for reference purposes at most as part of DDL and offered little in the description of its specific functions and effects on L2 writing.

Thus, this study is significant for several reasons. First, the study traced ESL graduate students’ independent uses of a suite of reference resources during an authentic writing task in a naturalistic setting, that is, a course paper or research proposal over the entire process of composing, ranging from planning to revising. Therefore, this study can contribute to closing the gap identified in the previous studies, which, to a large extent, examined only classroom settings or the editing/revising stages of the writing process. Second, unlike most previous studies, the present study examined how concordancing and other resources can complement each other (or not) by providing the participants with a suite of multiple resources that include not only different types of concordancing tools (i.e., two POS-tagged corpus tools and Google search engines) but also existing familiar reference resources that most of the participants had already been using (i.e., online monolingual and bilingual dictionaries and a thesaurus). These two considerations in terms of task and tool enhance the ecological validity of the study and help produce a fuller and more truthful picture of L2 writers’ reference resource consultation. Third, this study employed multiple methods such as screen recordings, a computer-generated
query log, and stimulated recalls to cause as little disruption for participants as possible and to generate, through triangulation, more accurate and reliable data on participants’ cognitive processes that occurred during interactions with the reference suite than had previous studies. Last, although corpus consultation was considered problem-solving activity in many previous studies, this present study is to my knowledge the first study that provides detailed and systematic descriptions and analyses of reference resource consultation during writing using the conceptual/analytic frameworks of problem solving established in cognitive psychology.
CHAPTER TWO
REVIEW OF RELEVANT THEORIES AND RESEARCH

This chapter discusses the literature relevant to L2 writers’ use of concordancing and other reference resources as a problem solving tool for their academic writing. It begins with an overview of relevant theories on two dimensions: (a) a brief discussion of major contributions of corpus linguistics to language pedagogy, and (b) specific theories underpinning the use of corpus tools and broader technology for L2 writing. Then follows a review of previous empirical studies into the use of corpus tools and Google, directly and indirectly, as a reference tool for academic writing. Finally, the types of problems arising during composition and those addressed by reference tools are discussed.

2.1 Corpus linguistics and its contributions to L2 pedagogy

Many of the potential benefits of using concordancing for L2 writing briefly introduced in Chapter 1 are to a large extent based on the major insights from corpus analyses that have informed L2 pedagogy as a whole. Drawing on naturally occurring linguistic data, corpus linguistics has been making a major impact particularly on the description of language in use, bringing to light linguistic aspects that are not obvious or readily accessible with elicitation- or intuition-based approaches (Cook, 1998; O'Keeffe, McCarthy, & Carter, 2007).

One major strength of corpus linguistics is that it provides an accurate and unbiased description of language as it looks at language in use, written and spoken by language users in naturally occurring settings (Biber, Conrad, & Reppen, 1998). Corpus linguistics is a significant departure from the traditional description of language by generative grammarians, who typically pursue theoretical discussions of idealized language that, to a large extent, intend to advance the knowledge of Universal Grammar, or “what humans know about language in advance of experience” (Meyer, 2002, p. 3). Linguistic data used in such inquiry mostly derive from linguists’ intuitions and elicitations from language users in laboratory settings (de Beaugrande, 2000; Meyer, 2002), which represents what Chomsky refers to as competence. By contrast,
corpus linguistics, by looking into attested, and authentic language data, demonstrate the performance aspect of language use.

This performance side of language foregrounded by corpus analyses has a number of major implications for language pedagogy. First, from the perspective of the communicative approach to language teaching, corpora are a great source for authentic texts. Real and authentic language provided by corpora can replace invented, contrived, or simplified language examples presented in traditional teaching materials, so that learners can see how language is actually used to communicate in context along with its complexities and nuances (Aston, 2000; McEnery & Xiao, 2011; McGarrell, 2013; O'Keeffe, et al., 2007).

Second, frequency information from corpus-based analyses shed light on the most common or typical patterns in use in a given context. Non-corpus based language textbooks also provide common and typical patterns, but these are only based on authors’ intuitions. Research conducted with large and balanced corpora (e.g., Biber & Conrad, 2001; Biber & Reppen, 2002; Mindt, 1996) has demonstrated that frequent patterns identified by corpus analyses are oftentimes inadequately accounted for or completely missing in syllabuses and teaching materials. Given the importance of frequency in L2 learning (for a comprehensive overview of the impact of frequency on L2 learning, see N. Ellis, 2002), these researchers suggest that frequency data provided by corpus analyses should be among the major factors that determine what to teach or what to include in syllabuses and teaching materials.

Third, corpus findings bring into sharp focus lexico-grammatical patterns in language. Language descriptions in formalist linguistics are based on the notion of the separation of grammar and lexis, in which sentences are constructed by filling individual words in slots categorized by syntactic rules. However, corpus analyses have showed that there are many semi-prefabricated phrases functioning as single units of meaning and that syntactic structures and lexis constrain each other and therefore cannot be easily separated (Sinclair, 1991). At the heart of lexico-grammar are collocations, which refer to combinations or sequences of words that co-occur more often than would be expected by chance. Corpus findings regarding collocation highlight the conventional and idiomatic aspects of language use (as opposed to the creative and generative ones) (Howarth, 1996; H. Yoon, 2005) and reveal that collocations are a central means by which to create meaning, and achieve fluency and naturalness in language use (Barfield, 2013; Webb & Kagimoto, 2011).
Lastly, corpus analyses based on specialized corpora have shown how register varies depending on the context of language use (Keck, 2012). For example, corpus-based research has provided detailed descriptions of the linguistic features of academic discourse, and analyses of linguistic variation across academic disciplines and genres (Biber, Johansson, Conrad, & Finegan, 1999).

2.2 Theories underpinning learner concordancing for L2 writing

As a theoretical basis, learner concordancing as a pedagogical tool has typically been associated with data-driven learning (DDL) (Johns, 1988, 1991), or discovery learning (Bernardini, 2004), which, in its widely practiced form, entails learners taking the role of researchers to work out rules and regularities out of concordance data on their own. More recently, in the broader context of technology use in education, concordancing and corpus technology have been seen as cognitive tools (or cognitive artifacts) that enhance and extend the learner’s cognitive abilities (Rüschoff, 2003; Park, 2010). These concepts, DDL and cognitive tool, in turn commonly draw on (cognitive and social) constructivism as their underlying theoretical foundation. In what follows, the constructivist learning theory, from which each concept derives, is briefly discussed. This discussion is then followed by a brief discussion of DDL and its limitations. Finally, the concepts of cognitive tools, distributed cognition, and affordances are provided as a possible theoretical framework from which to examine L2 writers’ corpus and online resource consultation for reference purposes.

2.2.1 Constructivist learning theory

Although diverse learning theories exist within the constructivist framework, at the core of these theories is the notion that meaning and knowledge are constructed and reconstructed by learners themselves, not transmitted intact from a teacher or outside sources. The notion is based on the epistemological assumptions that we individuals construct our personal realities in our minds by interpreting input through the filters of our previous experiences and therefore have different and unique representations of the world. This assumption stands in sharp contrast to the objectivist assumptions that objective realities exist and these realities can be mapped onto
the individuals and that knowledge is external to learners and thus can be transmitted (Jonassen, Davison, Campbell & Haag, 1995).

In the constructivist paradigm, learning is a process of constructing knowledge and making meaning, not a matter of acquiring the objective knowledge and demonstrating mastery of it (Duffy & Jonassen, 1992; O’Sullivan, 2007). Thus, learners are not seen as just recipients of external knowledge who only reproduce it, but rather active constructors of knowledge who are “engaging, grappling, and seeking to make sense of” their experiences (Perkins, 1992, p. 49). Conversely, it is argued that learners learn and understand best when they actively construct their own knowledge and create meaning in the process of solving problems or completing tasks that allow them to make connections to their lives. Knowledge constructed in this way is not inert but can be reconstructed and adapted to new and different situations. This aspect is particularly important in the information age, where the conception of knowledge itself has changed from possession of pre-constructed facts to the ability to make meaning out of novel information and to use it to solve problems at hand (Cobb, 1999; Jonassen & Reeves, 1996).

Constructivist education assigns the teacher the role of a facilitator—as opposed to the sole purveyor of knowledge in the objectivist paradigm—who monitors the process of personal meaning making, creating learning environments that encourage learners to examine their thinking and learning processes (Jonassen et al., 1995). Specifically, these constructivist learning environments are ones that help learners to learn how to learn by fostering learners’ mental activity, cognitive and metacognitive knowledge and skills, learner autonomy, and self-regulation of learning (O’Sullivan, 2007).

2.2.2 Data-driven learning (DDL)

Tim Johns is generally known to have used the term data-driven learning (DDL) for the first time in conjunction with corpus-based language learning (Boulton, 2011). Although Johns never gave a clear and distinct definition to DDL, the term generally refers to a language teaching and learning approach that involves the hands-on use of corpora by learners for language learning purposes. In typical DDL activities, learners explore multiple samples of the target language from corpora to discover patterns and rules on their own inductively instead of being explicitly taught about them. Johns described DDL as “the attempt to cut out the
middleman as far as possible and to give the learner direct access to the data” (1991, p. 30) and used the widely-quoted metaphors of “the learner as ‘linguistic researcher,’ testing and revising hypotheses, or as a ‘language detective,’ learning to recognise and interpret clues from context (‘Every student a Sherlock Holmes’)” (Johns, 1997, p. 101). Thus, DDL can be seen in close alignment with the constructivist paradigm in that learners are encouraged to explore language data to construct their own knowledge on the language they are learning.

Indeed, researchers on corpus-based language pedagogy argue that DDL fits well with the constructivist principles of language learning as it helps language learners develop cognitive and metacognitive skills and facilitate autonomy (Bloch, 2007; Boulton, 2010a; O’Sullivan, 2007). When formulating queries, sorting through and analyzing language samples from a corpus, learners engage in various types of cognitive activities such as noticing, hypothesizing, making inferences, and verifying. While doing so, learners become aware of how they go through the process of concordancing, or reflect on their thinking and learning processes and thereby further develop metacognitive strategies to improve them. The process of selecting areas to explore and techniques to use, formulating their own questions and drawing their own conclusions enables learners to take greater control over and responsibility for their own learning, thereby fostering autonomy in learning.

In practice, however, the majority of the previous empirical studies on corpus-based language learning treat DDL simply as a new instructional technique to use to teach given lexical and grammatical items (e.g., logical connectors) inductively (Boulton, 2011). Only a small number of studies have taken DDL activities further to let students choose their own language topics for exploration and formulate rules about them on their own (Boulton, 2010b). The latter type of DDL activities (more faithful to constructivist principles) has only been conducted in corpus linguistics courses or EAP courses with a focus on the use of corpus technology, where learners were generally assumed to have similar linguistic interests and competences to those of teachers or linguists (Bernardini, 2004) and a substantive amount of time and attention was allowed for DDL.

Thus, these widely practiced forms of DDL (as an approach for inductively learning linguistic items or a linguistic research tool for deriving general rules) discussed above may be limited in informing the use of concordancing for reference purposes or as an immediate writing support by L2 writers to solve problems they encounter in more typical settings of their
academic writing. Within the constructivist framework, a broader theoretical basis for problem solving through concordancing can be found in the concepts of cognitive tools and distributed cognition.

2.2.3 Cognitive tools and distributed cognition

Jonassen and Reeves (1996) argued that one way of creating a constructivist learning environment is the use of cognitive tools afforded by new technologies. In his earlier publication, Jonassen (1992) defined cognitive tools as technologies “that support cognitive processes, those that enable learners to engage in higher order thinking, […] , or that allow learners to generate and test hypotheses in meaningful problem-solving situations” (p. 6). In other words, cognitive tools amplify, extend, and reorganize users’ cognitive functioning and thereby facilitate their knowledge construction and meaning making. Technology use in education (e.g., universal computers, Webcasts, and spreadsheets) in the transmission paradigm is simply treated as a new and more effective means for transmitting and acquiring prescribed knowledge. Within the constructivist environment, the technologies may serve as cognitive tools that enable learners to access information, analyze the world, and interpret their experiences on their own. Jonassen and Reeves (1996) gave an example of spreadsheet software (e.g., Microsoft Excel) being used as a cognitive tool: using it to answer a “what if” question, e.g., “what if interest rates increase 1%?” A change in one cell will automatically recalculate related values in other cells and present results. The spreadsheets amplify and reorganize the user’s mental functioning, enabling her to engage in higher-order thinking (hypothesis testing) rather than a series of calculations. However, this process does not mean that cognitive tools possess inherent cognitive powers. It is still the learner that engages in noticing, inducing, and drawing conclusions (Wolff, 1997). In other words, “cognitive tools enable mindful, challenging learning, rather than the effortless learning”, often implied by the use of technology in education (Jonassen & Reeves, 1996, p. 698). As with broader constructivist learning theory, a particular emphasis is placed on the context in which cognitive tools are used. If cognitive tools are used for completing a task or to solve a problem that is meaningful to learners, it will result in greater learning (Jonassen & Reeves, 1996; Rüschoff, 2003).
These characteristics of cognitive tools can also be ascribed to concordancing tools (Rüschoff, 2003; Wolff, 1997). Concordancers sort and display language data in ways that allow users to discover patterns, test hypotheses, or figure out solutions to language problems at hand. These higher-order cognitive activities facilitated by concordancers would not be possible or would take much more time and cognitive effort when done by the user alone. Yet, concordancers do not analyze language data. It is the user who does all the decision making and problem solving at every stage of her analysis (Bernardini, 2004) with the concordancing tool helping her to overcome the limitations of her mind on memory and computing. This intellectual partnership between the user’s mind and cognitive tools (Salomon, Perkins, & Globerson, 1991) can be seen as a cognitive division of labor (Park, 2010) between the human mind and cognitive artifacts in distributed cognition, a sub-discipline in psychology.

Distributed cognition (DC) is grounded in the notion that cognition is shared and co-constructed in a social system across internal human minds and external cognitive artifacts within the system (Hutchins, 1995; Zhang & Patel, 2006). This shared cognition can materialize itself in a division of labor, in which different members pool and coordinate their resources to meet their needs. These different members here include cognitive artifacts. Cognitive artifacts—a concept in DC that is almost equivalent to cognitive tools discussed above—are tangible and intangible devices that maximize the cognition of the user by maintaining, displaying, and operating upon information (Norman, 1993). In a cognitive division of labor, these artifacts participate to take some of the cognitive load off the user by carrying out repetitive and time-consuming tasks (e.g., performing calculations, storing and retrieving data), and in turn allowing the user to focus on higher-order cognition (e.g., hypothesizing, inducing).

When this conceptualization is applied to the present study, the intellectual partnership between the L2 writer and the corpus tool can occur in the following manner: the tool carries out repetitive and tedious tasks such as sorting through data and displaying search terms and their contexts in a format in which patterns can be noticed effectively, helping the user to focus on higher-order cognition evaluating and recontextualizing patterns she notices from the concordance for her own writing (Park, 2010). Interestingly, corpus consultation features another layer of cognition of a distributed nature. A concordancer not only participates in the cognitive processing along with the user, but the concordance lines it presents are, seen
collectively, linguistic intuitions or knowledge shared and distributed across the texts compiled together in the corpus.

DC researchers suggest that what we should attend to when looking at a cognitive division of labor is how the human mind and the cognitive artifact interact to produce a synergistic outcome (or lack thereof). This is because the properties of a distributed cognitive system can be drastically different than those of each component. That is, the properties of the intellectual partnership cannot be inferred from the properties of the tool or the unaided human mind alone (Norman, 1993; Zhang & Patel, 2006). Interactional relationships of cognitive artifacts and their users are in turn affected by many factors such as the nature of a task, the level of expertise and skills of the user, and the socio-cultural context in which the user is situated (Norman, 1993). In this regard, affordance may be one theoretical concept through which to look at differences in individual users’ views and uses of cognitive tools.

2.2.4 Affordance

Affordance is a key construct in the ecological theory of language learning (van Lier, 2000). It is defined as “a particular property of the environment that is relevant—for good or ill—to an active, perceiving organism in that environment” (Gibson, cited in van Lier, 2000, p. 305). The environment provides different properties that are potential affordances for an organism. Active and engaging organisms perceive different affordances depending on what they want and what they think are useful. Applied to the language learning environment, an affordance is what provides language learners with “meaningful actions and interactions” in a given environment and what is meaningful can vary depending on the individual learners, that is, how they “perceive and relate to their immediate learning environment, including the tasks and tools with which they are engaged” (Hafner & Candlin, 2007, p. 305). Investigating the use of a corpus tool by students of law, Hafner and Candlin (2007) used the concept of affordance. They specifically examined whether and how L2 writers in a specific domain perceived affordances from the corpus tool, and then used them to support their legal writing. Results showed that the affordances the participants drew from the corpus tool were affected by their writing practice and culture in their discipline and their identity as a member of the disciplinary community. In light of the present study, this finding suggests that despite the potential benefits of a corpus tool
generally described above, individual users may perceive and adopt affordances from different aspects and features of the concordancing tool depending on factors such as the nature of tasks (whether authentic and thus meaningful or not), disciplinary culture, perceived value of writing, learning style, and identity as a L2 learner and writer. Individual differences, if any, will result from the complex interaction of these factors.

2.3 Empirical studies of concordancing as a reference resource

In this section, I review empirical studies that examined the use of concordancing as a reference tool by L2 writers. The review is broken down into three sections: (a) studies that examined participants’ performance in one-time corpus tool use for written error correction or revision tasks, (b) studies that traced participants’ longer-term use of concordancing and other reference resources during academic writing, and (c) studies that examined the use of Google in L2 writing.

2.3.1 Concordancing for error correction and revision

The majority of the studies that examined L2 writers’ use of concordancing as a reference source have involved one-time classroom tasks after a period of training, in which participants consulted corpora while correcting errors in their written productions or revising them based on teacher/researcher feedback or on their own.

Chambers and O’Sullivan (Chambers & O’Sullivan, 2004; O’Sullivan & Chambers, 2006) conducted a two-phase research project involving native English speaking learners of French at different levels of writing proficiency in Ireland. In the first phase (Chambers & O’Sullivan, 2004), eight graduate students of French wrote a 600-word commentary in French on a text relating to an aspect of the French language. The essays were marked with errors underlined. After training in concordancing over a three-week period, the students corrected the errors they chose through corpus consultation. The results were positive overall. Out of 85 changes made, 64 changes or about 75% were made correctly. In the second phase (O’Sullivan & Chambers, 2006), 14 undergraduate French major students followed the same procedure as
their graduate counterparts. Out of 166 changes the undergraduate participants made through corpus consultation, 122 changes (about 73%) were made correctly. Corpus use proved to be particularly useful for both groups in learning lexico-grammatical patterns and in reducing native language interference. In relation to prepositions and idiomatic expressions, many positive changes were made that would not have occurred as a result of consulting dictionaries and grammar books. When asked about how helpful working with the corpus was to improve their writing skills in French and whether they would keep using corpora, the postgraduate students were slightly more positive than the undergraduates on both questions.

In an academic writing course in a Japanese University, Gilmore (2009) conducted a similar study to examine the effects of training learners in the use of online corpora on the ‘naturalness’ of their revised essays. Forty-five Japanese undergraduate students enrolled in the course went through the same procedure as the participants of Chambers and O’Sullivan’s study, but the differences were that the participants in Gilmore’s study used two freely available online corpora (the Collins COBUILD Corpus and the British National Corpus) and that the native English-speaker raters compared the revised sentences with the original ones identified as problematic and determined which sounded more natural instead of rating the accuracy of the corrections. Out of a total of 350 problems identified in the first drafts, 214 or about 61% were rated as more natural. This suggests, the researcher argued, that online corpora can play a valuable role in helping students improve the naturalness of their writing.

How user training would make differences in the processes and outcomes of learners’ corpus use is one major theme commonly discussed in concordancing studies. Kennedy and Miceli (2001) offered a sense of how a gradual and guided training in learner corpus consultation can be done as part of a writing course. The researchers adopted an “apprenticeship” approach to training in order to “promote learning by example and by experience” (p. 79). The students practiced using a corpus specially designed by the researchers as a problem-solving tool when revising their own writing and also as a “treasure-hunting” tool (p. 79) to find typical expressions or patterns used for specific situations. To evaluate how successful the students would be at extracting information they need from the corpus without the help of a teacher, the researchers gave the students two texts to revise using the corpus. The results revealed that although the students made many successful investigations, they were ignorant of “common pitfalls and techniques for avoiding them” (p. 81). The authors concluded that their
apprenticeship training was not sufficient to equip the students with needed corpus investigation skills. They proposed, as a solution, training that focuses more on improving students’ observation and logical reasoning.

Building on the insights from the study, Kennedy and Miceli (2010) further developed their apprenticeship approach to training in corpus use and evaluated it through case studies of three individual students. In their new apprenticeship approach, they focused on corpus consultation directly related to the specific requirements of a writing task at hand while deemphasizing the learner-as-researcher approach. The two specific functions of corpus consultation they introduced were (a) pattern hunting aimed at getting “ideas on what to write about and words to express it” (p. 32), and (b) pattern defining aimed at finding models for a specific target pattern in terms of component words and structure. The researchers conducted this new apprenticeship in their Italian language course over a semester. To evaluate the new approach, at the end of the course, three students were asked to enrich their autobiographic writing they had written earlier in the course for 45 minutes using the corpus and other reference resources. Results showed that the individual learners’ attitudes towards and understanding of corpus consultation and their computer/Web literacy influenced the extent, purpose, and outcome of their corpus use. While two students used the two functions of corpus consultation effectively and one of them even developed a new function that was not taught (i.e., expanding on dictionary information through corpus investigation), one student showed limited use of the corpus, who failed to see pattern hunting as a unique function of corpus consultation that set it apart from other reference resources. Based on the results, the researchers drew insights for future corpus use training in which the unique functions of corpus consultation, which give it advantages over other resources, should be taught more explicitly. The researchers also suggested that a future apprenticeship approach include training in not only actual skills required to conduct searches and interpret examples but also appreciation of the nature of corpus consultation, such as a long process of trial and error to develop the proper skills.

Finally, although the task carried out by the participants was not a composition task, Frankenberg-Garcia’s (2005) study is worth a close look in that it was virtually the only study that directly investigated the reference resources the participants preferred depending on the particular needs for consultation, and compared the effectiveness of the resources and the participants’ perceptions of their effectiveness. Sixteen undergraduate student translators at a
Portuguese university translated a short newspaper article (about 230 words) from Portuguese into English, consulting any resources they wished to use among more than 20 resources provided (paper and online dictionaries, online concordancing programs, a term bank, and search engines). While translating, the participants kept a search log recording in detail the problems they encountered and the resources they consulted to solve them. Results showed that while 65% of the total look-ups were successful, the participants perceived 83% to be successful revealing a gap between actual and perceived effectiveness of the resources. Analysis of the participants’ logs revealed that while the most frequent purpose for which the resources were consulted was to find an L2 equivalent (58%), which was followed by confirming a hunch (18%), and finding a collocate (16%), the participants preferred resources mediated by linguists, and lexicographers (i.e., term bank, and dictionaries) by far to those requiring a greater amount of cognitive effort for interpretation (corpora and search engines). One major trend observed was that the participants tended to consult a single resource instead of moving back and forth between different resources to solve individual problems. Although the task in the study was somewhat different from those of other studies, the findings from this study provide two major insights. First, each resource may have unique functions for which it is best suited, and therefore resources may complement each other when used together. Second and correspondingly, learners should be trained not only in individual resources but also in how to consult them in combination.

The studies reviewed in this section provide a sense of how and to what extent concordancing can support L2 writers’ problem solving especially at later stages of the writing process. However, the task or research settings of these studies are limited in two aspects, such that the findings from them may not be generalized to more typical and common academic contexts. Firstly, the majority of the studies reviewed here used a small specialized corpus, which was not publically available, developed for the target course (Chambers & O’Sullivan, 2004; Kennedy & Miceli, 2001, 2010; O’Sullivan & Chambers, 2006). More importantly, these studies more or less looked at only the results of the participants’ one-time uses of corpora and other reference resources while performing mostly timed writing tasks in language/translation classroom settings. These task settings may be too limited to provide general ideas of how corpus tools can be used as a reference tool during the entire writing process and in more typical and authentic writing contexts for L2 writers in tertiary education. More specifically, timed
writing tasks used in these studies do not represent the typical academic writing tasks in which university students engage. For example, a term paper in a content course, one of the most commonly performed writing tasks in university settings, may take an extended period of time to complete and go through different, often recursive stages. Therefore, the findings from these studies offer little information about possible differences in corpus consultation patterns between different stages of the writing process or possible changes that come with corpus tool use in participants’ writing performance, and their attitudes and approach to academic writing.

2.3.2 Longitudinal concordancing studies

The studies to be reviewed here investigated over the course of a semester or longer how university students used corpus tools for their authentic writing assignments and how the corpus use affected their writing processes and their attitudes toward English academic writing. These studies traced, to varying extents, participants’ independent corpus use over time while completing their writing assignments, using methods such as search logs and/or screen recording.

Hyunsook Yoon (2005, 2008) conducted a two-phase longitudinal case study on ESL graduate students taking an EAP course. The EAP course had corpus consultation as a regular activity throughout the semester. The researcher examined, in the first phase, how corpus use (Collins COBUILD Corpus) changed the L2 writing process and competence of six focal students during the semester. For the second phase, she investigated how and to what extent the same students used the corpus independently in the subsequent semester. One major finding was that although most students used the corpus as a reference tool for problem solving purposes, corpus consultation also helped the students to raise their language awareness, especially of the importance of common usage and collocation, which in turn affected their approaches to writing and the writing process. By checking problems with the corpus while writing, they took more responsibility for their writing and became more independent learners. Another finding of note was that participants’ fields of study rather than their initial attitude toward the corpus tool seemed to play a greater role in determining the frequency and scope of corpus use. Participants who were science majors (some of whom showed the most positive attitudes toward corpus use) had fewer writing assignments than the non-science major participants and used the corpus less
frequently. Yoon argued that this need-based corpus use confirmed the findings from broader CALL studies that “lack of meaningful engagement” (p. 44) with the technology would undermine its appeal and learners’ motivation to use it. Based on the findings, the researcher concluded that even advanced-level L2 writers who are well acculturated into disciplinary writing in terms of content and idea development can still struggle with linguistic features and that corpora can in that regard serve as a significant instructional and learning tool.

Unlike other studies on concordancing, which were conducted mostly in general language or academic writing courses, Hafner and Candlin (2007) examined the potential of corpus use as an affordance for students of law in terms of providing language support for their legal writing assignments. Two cohorts of law students at a Hong Kong university were asked to consult a specially compiled online corpus of legal cases as a language reference tool when completing their legal writing assignments. Results showed that the participants’ independent corpus use was modest. About 30% of the participants used it without prompting for their own purposes. The authors attributed the relatively low rate to the participants’ unfamiliarity with the corpus tool and their use of other resources such as Google for language support. Interestingly, analysis of the search terms used by the participants revealed that their searches were more intended to get full exemplar legal documents than knowledge about specific lexical items. Linking this to professional practices, where legal documents are drafted following legal precedents, the authors concluded that writers may perceive and adopt different affordances from corpus tools depending on their actual needs and the disciplinary/professional practices they are exposed to.

These two studies suggest that motivation, extent, and success in using a corpus tool as writing assistance can vary depending on the needs of individual learners, their familiarity with the tool, and affordances they perceive to be relevant to their tasks.

Finally, in a study that offers especially detailed descriptions of the functions of corpus consultation, Park (2010) examined the potential of a domain-specific online corpus and Google-powered concordancer to improve the lexico-grammatical awareness and performance of ESL undergraduate students enrolled in a genre-based academic writing course. Triangulation of data collected from an automated query log, screen recordings and stimulated recalls over a semester allowed the researcher to produce detailed descriptions of functions of corpus consultation and participants’ search patterns. Overall findings suggested that the corpus tool
can be an effective and efficient tool in improving students’ lexico-grammatical awareness of a
given genre and their performance in its textual realization. However, the author also found that
the students did not use the corpus consultation to its full potential. For example, they consulted
the corpus mostly to confirm the accuracy and appropriateness of language items rather than to
elicit exemplar academic texts. Based on the patterns and functions he identified from the
students’ searches, and analyses of interactions between the students and the corpus tool, the
author concluded that corpus consultation should not be seen as a tool for DDL but a cognitive
artifact that participates in collaborative co-construction of text with its human user.

These studies made a major contribution to the research of learner concordancing by
providing detailed descriptions of learners’ interactions with the corpus tools and tracing the
changes brought about by corpus tool use in various aspects of their academic writing. However,
as with the studies reviewed in the previous section, these studies are not without limitations.
Chief among them is the fact that they did not reflect the real-life choices that today’s L2 writers
have with respect to reference resources (Frankenberg-Garcia, 2005) by providing the corpus
tools as the only or main resource to consult during the tasks. Although there were a variety of
reference resources available to use, some of which the participants had already been using,
these studies neither offered the participants any training in combined uses of different reference
tools nor did they examine such uses.

2.3.3 Google as writing assistance

On the strength of its abundant authentic and up-to-date language data, the Web has been
increasingly used as a corpus both for linguistic research and language teaching and learning.
Powerful search engines, represented by Google in particular, have been suggested as an easy-to-use concordancer with which learners can sift through the abundance of data provided by the
Web to get immediate language support and help learn their target language (Chinnery, 2009;
Sha, 2008; Shei, 2010).

Recently, researchers have been starting to go beyond suggesting ways of practicing
what has been referred to as Google-assisted language learning (GALL) to empirically examine
the effects of the use of Google on L2 learners’ writing (Conroy, 2010; Fujii, 2007; Park, 2010).
In what follows, I will briefly review studies that explored the potential of Google for
supporting L2 writers. Park (2010), which used Custom Search Engine (CSE), a customizable search engine provided by Google, is excluded because it has already been reviewed in the previous section.

Motivated by a lack of writing support for English as an additional language (EAL) students at Australian universities, Conroy (2010) investigated how Internet-based reference tools would help EAL students improve their academic writing and become more independent language learners and L2 writers. After a brief training on how to use corpus sites, Google, and online dictionaries for their English writing, 165 EAL university students were given 2 to 4 weeks to use the tools on their own when writing. They were then surveyed on their uses of and attitudes towards these tools before and after the training and practice. The vast majority (about 90%) of the students showed positive attitudes toward GALL and concordancing for language learning and writing support. Of particular note was that while many of the students had already been using Google as a reference tool even before the training, their initial search techniques were simple and crude, mostly intended for content searches. Specific search techniques and strategies—needed to use Google as a concordancer—covered in the training were novel to the students and therefore perceived as very useful in supporting their academic writing. Another interesting finding was that the students preferred Google to the online corpora. Based on these findings, the author suggested that Australian universities with limited writing support staff provide EAL students with training in the use of GALL and concordancing techniques as a way to independently monitor and improve their writing.

In another setting where language support for learners is not so readily available, Fujii (2007) examined the potential of Google for helping learners improve the accuracy and naturalness of their translation from L1 to L2. After studying a handout on how to effectively retrieve words and phrases from Google and integrate them into their own writing, 176 science and engineering majors in a Japanese university were asked to rewrite abstracts they had written earlier consulting only dictionaries. They were also surveyed on their attitudes toward the use of Google for their translation work. Although Fujii provided no specific quantified measures on participants’ improvements, the author presented overall positive findings: the students showed improvements in several areas such as articles, singular/plural distinction, and verb tenses, which are traditionally problem areas for Japanese learners of English. The survey results also
revealed that the students were enthusiastic about the innovative ways to use Google for improving their writing.

Watson Todd (2001) also reported overall positive results in a study where 25 Thai post-graduate students in an English language support course induced rules from self-selected concordances and applied those rules to self-correction of the errors in their writing. The participants’ writing samples were marked by the researcher for lexical items that were misused. The participants then chose one of the marked lexical errors, searched for instances of the lexical item on the Internet, selected 10 concordance lines, extracted the rules/patterns from those 10 concordances and corrected the lexical errors they had made applying those rules. The results were that a mean of 7.78 concordance incidences (out of 10) matched the induced patterns, and that out of a total of 23 lexical items, 18 were validly corrected. One major factor that helped to produce the overall positive results here was that the participants selected only 10 incidences that they could comprehend and therefore could avoid “overexposure to concordance lines” from the Internet (Thurstun & Candlin, 1998, p. 278). In addition, they were given opportunities to make immediate use of induced knowledge by applying it back to the correction of their errors.

These studies present overall positive findings on GALL application in L2 writing particularly as a handy reference tool. However, researchers also point out that GALL comes with its own limitations (Conroy, 2010; Fujii, 2007; Wu, Franken, & Witten, 2009). First of all, the abundance of data can overwhelm learners, especially students with low language proficiency, in terms of the time it takes to sort through data and the cognitive load of interpreting the search results. Secondly, Web content, by nature, is extremely heterogeneous, a mishmash of different varieties, genres, and registers (Wu et al., 2009) and therefore poses a challenge for learners in that they have to figure out contexts from the search results. Thirdly, the Web does contain a great deal of lexical and grammatical errors by both L1 and L2 speakers, and therefore may often do a disservice to learners who turn to GALL for improving their language accuracy.

To ease, if not resolve entirely, these problems, researchers have been suggesting a number of strategies and alternatives: (a) checking basic information provided in results (frequency and domain) to evaluate the appropriacy of the target items, (b) using advanced search options—e.g., using a wildcard or search operators—to get more focused results (Conroy,
2010; Sha, 2010), (c) using specialized search engines such as Google Scholar and CSE (Brezina, 2012; Park 2010) to confine search results to a particular domain, and most importantly, (d) using Google along with other reference resources in complementary manners (Kaur & Hegelheimer, 2005; Pérez-Paredes et al., 2012). Nonetheless, despite these abundant and various suggestions, there has been little empirical research on whether and to what extent these strategies help L2 writers to overcome the limitations of Google for language support.

2.4 Concordancing as a problem solving tool

From the cognitive view of writing, composing has traditionally been conceptualized as a recursive, cognitively demanding, problem solving activity that involves continuous decision-making at multiple levels (Flower, 1989; Flower & Hayes, 1981; Heine, 2010; Manchón & Roca de Larios, 2007). Consulting corpora and other resources during composition can then be viewed as part of the writer’s problem solving effort to find immediate solutions to difficulties with text formulation and refinement in particular. Indeed, in many of the previous studies, corpus tools are presented as problem-solving tools for linguistic problems (Kennedy & Miceli, 2001, 2010; Park, 2010, 2012; Thurstun & Candlin, 1998; H. Yoon, 2005, 2008). In this regard, it is worthwhile to review (a) what types of problems that L2 writers encounter during text composition and (b) what types of problems they actually tackle using corpus tools.

2.4.1 Types of problems arising during text formulation

Various types of problems can occur during composition on multiple levels related to content, organization, audience expectations, and language use. However, given that reference resource consultation for writing is particularly concerned with language use, this section focuses only on the kinds of problems that commonly arise in the process of text formulation, specifically lexical searches and syntactic encoding.

In a study that explored the language learning potential of L2 writing, Cumming (1990) identified three types of form-meaning mapping problems, in which participants attended to metalinguistic and ideational concerns concurrently: (a) searching for and assessing appropriate words and phrases, (b) comparing cross-linguistic equivalents, and (c) reasoning about linguistic
choices. Similarly, Roca de Larios, Manchón, and Murphy (1996) classified the problems their study participants tackled in the process of text formulation into four types depending on the availability of the linguistic form to express a given intended meaning:

P1: The writer has constructed or is in the process of constructing a mental representation and has to retrieve elements from long-term memory to express it.
P2: The intended meaning is encoded in the L1 and tries to find an equivalent in the L2.
P3: The writer has an option available that expresses or conveys the already constructed mental representation (or intended meaning) but tries to upgrade it in conceptual or linguistic (stylistic) terms.
P4: The writer has an option available to express the intended meaning but has doubts as to its correctness or appropriacy. (For full descriptions, see ibid. p. 10)

In a series of follow-up studies looking into the problem solving nature of text generation by L2 writers (Manchón, Roca de Larios, & Murphy, 2009; Murphy & Roca de Larios, 2010; Roca de Larios, Murphy, & Manchón, 1999; Roca de Larios, Manchón, & Murphy, 2006), the researchers merged these four types into two broad ones: compensatory and upgrading problems. **Compensatory problems** refer to problems caused by lack of access to the linguistic knowledge needed to express the intended meaning while **upgrading problems** result from an effort to find a more appropriate expression or a better match for a given language item or idea (Roca de Larios et al., 2006).

Figure 2.1 juxtaposes the problem types discussed above according to their similarities. Taken together, form-meaning matching problems encountered by L2 writers during text formulation can be seen to involve varying degrees of two processes: **searching** for a (lexical and grammatical) form and **evaluating** a form against the given intended meaning.

<table>
<thead>
<tr>
<th>Roca de Larios et al.</th>
<th>Cumming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compensatory</td>
<td></td>
</tr>
<tr>
<td>P1: The writer has constructed or is in the process of constructing a mental representation and has to retrieve elements from LTM to express it.</td>
<td>Searching for and assessing appropriate words and phrases</td>
</tr>
<tr>
<td>P2: The intended meaning is encoded in the L1 and tries to find an equivalent in the L2.</td>
<td>Comparing cross-linguistic equivalents</td>
</tr>
<tr>
<td>P4: The writer has an option available to express the intended meaning but has doubts as to its correctness or appropriacy.</td>
<td>Searching for and assessing appropriate words and phrases</td>
</tr>
<tr>
<td>P3: The writer has an option available that expresses</td>
<td></td>
</tr>
</tbody>
</table>
### Upgrading

or conveys the already constructed mental representation (or intended meaning) but tries to upgrade it in conceptual or linguistic (stylistic) terms.

<table>
<thead>
<tr>
<th>Upgrading</th>
<th>Reasoning about linguistic choices</th>
</tr>
</thead>
</table>

*Figure 1.1. Types of language problems that arise during text formulation.*

#### 2.4.2 Types of problems addressed with corpus tools

One way of looking at the types of problems that can be tackled consulting corpora and other reference resources is through investigating the purposes of reference resource consultation identified in the previous studies. I define here the purposes of reference tool consultation as the needs that arise while writing and prompt the user to carry out a specific query (or look-up). Although there is a considerable body of research on learner concordancing, there are only a small number of studies that actually provide descriptions of specific query purposes at any level of detail. In these studies, query purposes are commonly expressed as functions (Frankenberg-Garcia, 2005; Kennedy & Miceli, 2010; Park, 2010).

First, in the study of student translators’ reference resource use, reviewed earlier, Frankenberg-Garcia (2005) classified her participants’ look-ups into five functions: (a) finding an L2 equivalent when identifying a corresponding L2 word or expression for the given L1 item; (b) confirming a hunch when learners try to check if a certain L2 form is acceptable in a given context; (c) finding a suitable collocate to identify which adjectives, nouns, or verbs are best combined with a given L2 word; (d) choosing the best alternative when determining the best choice out of multiple alternatives; and (e) checking spelling.

Second, Kennedy and Miceli (2010) trained their students in two broad functions of corpus consultation and also used them as codes for analyzing their case study participants’ queries. Pattern hunting is aimed at finding “ideas on what to write about and words to express it” (p. 32). Pattern defining, on the other hand, is aimed at finding models for a specific target pattern in terms of component words and structure. In addition, the researchers identified finding an L2 equivalent for a given L1 pattern as a typical function of the participants’ dictionary consultations.

Lastly, Park (2010) identified two broad functions from participants’ corpus queries. First, an elicitation query is used when retrieving accurate and appropriate words and
grammatical constructions that deliver a given intended meaning. Second, a *verification query* is performed when the writer has specific words or constructions in mind for a given intended meaning and confirms their accuracy or appropriacy through corpus consultation.

Figure 2.2 below juxtaposes the functions just discussed above according to their similarities. As can be seen in the Figure, these functions (or purposes) of reference resource consultation can be grouped into two broad types: *eliciting* new (lexical and grammatical) knowledge from the reference resources; and *confirming* the writer’s existing knowledge. However, this distinction is not clear cut as a reference resource consultation can have multiple functions and purposes (Park, 2010).

<table>
<thead>
<tr>
<th>Frankenberg-Garcia</th>
<th>Kennedy &amp; Miceli</th>
<th>Park</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finding an L2 equivalent</td>
<td>Finding an L2 equivalent to a given</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L1 pattern</td>
<td></td>
</tr>
<tr>
<td>Finding a suitable collocate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirming a hunch</td>
<td>Pattern hunting, Pattern defining</td>
<td></td>
</tr>
<tr>
<td>Choosing the best alternative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checking spelling</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2.2. Functions of reference resource consultation.*

To link these purposes of reference resource consultation back to the types of text formulation problems discussed in the previous section, one can see a considerable correspondence. To compare only at the broadest level, the two broad purposes, *elicitation* (of new knowledge) and *confirmation* (of existing knowledge), correspond to the two broad problem types, *searches* and *evaluation* respectively. This correspondence suggests that reference resources can be used to solve many types of the lexical and grammatical problems that arise during L2 text generation.
CHAPTER THREE

METHODS

The present study employed a mixed methods design consisting of two phases. The first phase was an exploratory phase focused on the training of participants and their hands-on practice in the use of a reference suite developed for the study. Specifically, the participants were given time to familiarize themselves with the reference suite and to take it up as a supporting tool for their English writing during the first phase. The second phase involved case studies that followed six participants during the entire process of completing an authentic writing assignment independently while consulting the reference suite.

3.1 Participants

In order to investigate how ESL students independently consult reference resources for their English academic writing outside of the classroom, I recruited Korean graduate students residing in Toronto, Canada, who were enrolled in writing-intensive programs and would have at least one major writing assignment to complete during the study period. The rationale for recruiting this specific group of participants was as follows: First, graduate students are more likely than undergraduates to have the levels of (meta-)cognitive skills and (meta-)linguistic knowledge required to formulate queries, evaluate search results, and apply them to their writing on their own. They are also assumed to be more willing and motivated to invest time and energy in improving their writing performance using a reference tool for their academic success and career advancement. Second, by sharing the same L1 (Korean) with the participants, I would be able to directly observe and analyze participants’ natural writing/linguistic problem solving, if any, that involves L1-L2, both textual and conceptual, translation. Third, recruiting students in writing-intensive disciplines was to ensure that participants would have sufficient writing opportunities to engage in concordancing in their naturalistic settings. Previous studies (e.g., H. Yoon, 2005) showed that participants from science and engineering disciplines may not use corpus tools as often as those from writing-intensive disciplines such as the humanities and social sciences, even though they highly value the utility of a corpus tool.
I searched student profiles publicly available on the websites of graduate programs based in Toronto and identified Korean students enrolled in humanities and social science disciplines. I sent all of them the information letter (Appendix A) via email. I also asked colleagues and friends to distribute the letter to potential participants. Initially, eight students volunteered to participate. However, one participant withdrew from the study due to her busy schedule. Another participated in the training sessions but was not able to continue her participation because of some compatibility issue between her laptop computer and the reference tool program that should be installed for the research. Table 3.1 presents profiles of the six participants who participated for the entire duration of the study.

Table 3.1

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Jae</th>
<th>Yumee</th>
<th>Jinho</th>
<th>Goeun</th>
<th>Shia</th>
<th>Ian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Age</td>
<td>37</td>
<td>28</td>
<td>24</td>
<td>29</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>Degree pursued, Field of study</td>
<td>PhD, Adult Education &amp; Human Resources</td>
<td>PhD, Educational Technology</td>
<td>MA, Information Studies</td>
<td>PhD, Language Assessment</td>
<td>MA, Social Work</td>
<td>PhD, Second Language Education</td>
</tr>
<tr>
<td>Years in Canada</td>
<td>3.3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.3</td>
<td>10.3</td>
<td>1.3</td>
<td>0.8</td>
<td>1.2</td>
</tr>
<tr>
<td>Education in Korea</td>
<td>BA in Law, MA in Adult Education</td>
<td>BA, MA in Educational Technology</td>
<td>Immigrated to Canada at grade 8</td>
<td>BA, MA in English Linguistics</td>
<td>BA in Social Work</td>
<td>BA, MA in English Education</td>
</tr>
<tr>
<td>Prior experience with corpus tools</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>COCA, and COBUILD</td>
<td>None</td>
</tr>
<tr>
<td>Approximate duration of participation</td>
<td>18 weeks</td>
<td>20 weeks</td>
<td>16 weeks</td>
<td>12 weeks</td>
<td>7 weeks</td>
<td>16 weeks</td>
</tr>
</tbody>
</table>

<sup>Note</sup>: a. All participants’ names are pseudonyms.
<sup>b</sup>. This period includes three years Jae studied in the U.S. before he came to Toronto.

Ages of the participants ranged from 24 to 37 years at the time of data collection. Genders were evenly distributed into three female and three male students. The participants were all from writing-intensive disciplines with four, all doctoral students, from the broad discipline of education, and two, MA students, from information science and social work
respectively. Most participants had no previous experience of living and studying in English speaking countries before they came to study in their current programs. The exception was Jinho, who immigrated to Canada with his family when he was 13 years old and finished his high school and undergraduate education in Canada. But he still considered Korean his first language and reported that he did L1-L2 translation frequently while writing. When it comes to prior experience with corpora and concordancing tools, most participants had none and had even never heard of the word “corpus” and its uses in language learning. Goeun was the only one who had been occasionally using COCA (described in detail in the following section) and the online sampler of COBUILD concordancer software.

The participants can be divided into two groups in terms of the timing of their participation. Jae, Yumee, Jinho and Goeun started their participation during the period between December 2011 and January 2012 in the space of a few weeks from one another. They all had a writing assignment to complete during the spring term. Shia and Ian, meanwhile, joined the study in May, 2012 and completed their writing assignment during the summer. Participation duration also varied, ranging from as short as seven weeks to 20 weeks at the longest. The wide variation in the participation period was caused mainly by different types of writing assignments the participants completed for the study and their varied time availability for the training and their individual assignments. It should be noted, however, the participation period here only means the time taken to finish the entire tasks (four screen recordings and stimulated recalls, see below), which were spaced depending on the individual participant’s schedule. The frequency and intensity of their reference tool uses during those periods also differed widely.

3.2 Reference suite (i-Conc)

With the help of a former colleague and computer programmer, I developed a mini web-browser program named i-Conc (short for Internet concordancer) for the study (see Figure 3.1). It allows the user to access multiple language reference resources that are freely available on the Web. The program features four tabs, in which different types of Web concordancers, and online dictionaries can be accessed.
Specifically, five concordancing resources and three types of dictionaries are provided in i-Conc. Table 3.2 presents these resources featured in each tab. There were two rationales for developing a referencing suite that features the eight different resources, which are free to use by visiting their individual sites. First, it would be very cumbersome if participants had to visit each individual site to consult the resources while writing, often ending up with multiple windows open on their computer screen. i-Conc allows the user to access these resources in a single interface making it easier to move back and forth between multiple resources on the same window. Second, to observe participants’ independent uses of a reference tool, it is important to trace searches that participants perform. Recording the content of searches manually whenever a participant consults a reference resource, as was done in some of previous studies (Frankenberg-Garcia, 2005; H. Yoon, 2005), would be time-consuming and might significantly interrupt the participant’s actual writing process. This complication may cause participants to avoid consulting reference resources altogether or use them less often than they would if they do not have to keep such a log. i-Conc has its own search box in addition to search boxes provided on each individual site (see Figure 3.1). Every query performed in the i-Conc search box is automatically recorded in a query log, another program developed for the study, (see Section

---

1 Queries entered in the main search box located in the top frame are automatically recorded, but queries done in the COCA tab are exceptions. As COCA’s interface has multiple search boxes and options that have to be selected for sophisticated searches, it is technically not possible to use an external search box (i.e., the main search box of i-Conc) to access the corpus. To record queries entered in COCA, therefore, the user has to use COCA’s own
3.3.2.4 Query log for detail). With the query log program running in the background, participants can by and large consult the tool without much interruption.

Table 3.2

Reference Resources Featured in i-Conc

<table>
<thead>
<tr>
<th>Resource type</th>
<th>Tab name</th>
<th>Resource name</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concordancer</td>
<td>COCA</td>
<td>Corpus of Contemporary American English</td>
<td><a href="http://corpus.byu.edu/coca">http://corpus.byu.edu/coca</a></td>
</tr>
<tr>
<td>Google</td>
<td>Google Web</td>
<td>Google Scholar</td>
<td><a href="http://scholar.google.com">http://scholar.google.com</a></td>
</tr>
<tr>
<td></td>
<td>Custom Search Engine</td>
<td></td>
<td><a href="http://www.google.com/cse">http://www.google.com/cse</a></td>
</tr>
<tr>
<td>JTW</td>
<td>Just The Word</td>
<td></td>
<td><a href="http://www.just-the-word.com">http://www.just-the-word.com</a></td>
</tr>
<tr>
<td>Dictionary</td>
<td>Dictionaries</td>
<td>Bilingual (Naver)</td>
<td><a href="http://endic.naver.com">http://endic.naver.com</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Longman Dictionary of Contemporary English</td>
<td><a href="http://ldoce.online.com">http://ldoce.online.com</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thesaurus</td>
<td><a href="http://www.thesaurus.com">http://www.thesaurus.com</a></td>
</tr>
</tbody>
</table>

In the following section, I go over briefly the features and characteristics of each resource presented in Table 3.2.

3.2.1 Corpus of Contemporary American English (COCA)

The Corpus of Contemporary American English (COCA), linked in the first tab of i-Conc, is one of the largest corpora that are freely accessible on the Internet, containing about 450 million words. Created by Mark Davies of Brigham Young University, it is a balanced corpus of American English, consisting of texts from different genres such as spoken, fiction, magazines, and academic journals that were produced from 1990 to the first half of 2012. COCA thus contains the latest texts and gets updated regularly (Davies, 2008), which is not the case with other large publicly available corpora such as the British National Corpus (BNC).

COCA comes with a web interface through which the corpus is accessed. In this study, COCA refers to both the corpus and its web interface. The interface features a variety of search options allowing the user to search for, among others, exact words, phrases, synonyms, syntactic
structures, and collocations in specific genres, and by lemma, part of speech, and wildcards, with the results displayed in order of frequency. Being the most sophisticated concordancer featured in i-Conc, COCA allows the user to customize their searches to meet their specific concordancing needs. However, it also means that the user needs to familiarize herself with COCA’s query syntax (see Appendix B: i-Conc tutorial for key query syntax).

COCA provides four different options for displaying query results. Among them, LIST and KWIC are the most relevant and useful options that the user can use for referencing purposes while writing in English. So as not to overwhelm the participants with too many features and options available in COCA they were encouraged to use these two display options while writing. In the LIST option (see Figure 3.2), users fill out the search box in the left frame of the interface and click the SEARCH button. Then it returns a frequency listing of the items that match the query condition in the upper right-hand frame. When the user clicks on any of the entries in the listing, concordance lines containing the searched-for item in the middle are displayed in the lower right-hand frame. The LIST option is particularly useful to search for commonly or typically used words and phrases in a given context or to see different alternatives, from which the user can choose the closest to her intended meaning.

In the KWIC option, queries are filled out the same way as in LIST, but COCA returns in the lower right-hand frame random 100 (or 200 if set so) instances of the searched-for item in

![Figure 3.2. Screenshot of LIST display on COCA.](image)
KWIC format, which is the traditional concordance view, where the searched-for item is displayed in the center surrounded by sortable contexts to the left and right. What sets the KWIC format in COCA apart from other concordance displays is that each word in the immediate context on both sides of the searched-for item is highlighted in a different color according to its part of speech (see Figure 3.3). This option is useful especially for quickly checking typical and common syntactic patterns in which a linguistic item occurs, or its colligational relations.

![Screenshot of KWIC display on COCA.](image)

**Figure 3.3.** Screenshot of KWIC display on COCA.

### 3.2.2 Google

The second tab features three Google search engines: Google Web, Google Scholar, and Custom Search Engine (CSE) (see Figure 3.4). The Web is not considered a corpus in the strict sense as the texts within it are not collected in a principled way nor are they representative of a particular language variety (Biber, Conrad, & Reppen, 1998; Boulton, 2011). Commercial search engines such as Google are also different from traditional concordancers particularly in two aspects. First, search results are not displayed in KWIC format in which the searched-for items are centrally aligned. Second, the frequency information provided in search engines is not the number of tokens of the linguistic item searched for but the number of web pages that contain it.
Nonetheless, for the reasons provided by GALL researchers (as briefly discussed in Chapter 1), Google Web was included in *i-Conc* as a concordancer in this study. To briefly recap the advantages of using the Web as a corpus and Google as a matching concordancer, firstly, the Web with its enormous linguistic data can provide much richer results even for relatively infrequent words, phrases, and even long strings of texts than can any standard large-scale corpora. Secondly, Google is easy to use as it requires no complicated query syntax, and processes queries very quickly, which is an important factor for any online tools. Students are already familiar with it from years of use for their content searches and may already have developed their own strategies for effective searches. Thirdly, Google can offer unique ways of supporting L2 writers: Google search results are typically displayed in a “snippet” format, consisting of the title of a web page that contains the search item and two to three lines of context surrounding the item on that page (see Figure 3.4). Although not as effective as the KWIC format in helping the user notice common patterns across concordance lines, it gives broader contexts in which the target item occurs than does KWIC (single-line concordances) and provides direct access to the original source from the snippet comes from. Moreover, unlike most concordancers, which return concordances that exactly match the query, Google returns results that are close or similar to the query item as well as exact matches, allowing the user to explore the alternatives. As such, Google Web’s search engine can be used when the user needs...
to quickly check the existence and overall frequency of a linguistic item (especially multi-word strings) or elicit its collocational or colligational information in ways that are difficult or not possible with more traditional concordancers. Google Web can therefore be a useful complement to other more traditional concordancers featured in i-Conc.

However, as pointed out in Chapter 2, using Google as a concordancer is not without limitations. To give a short summary of them, Web content, by nature, is extremely heterogeneous and can be a mishmash of different genres, registers, and linguistic varieties, which can pose a problem for learners who intend to use it as a supporting tool for their academic writing. In addition, the Web contains a significant number of lexical and grammatical errors made by both L1 and L2 speakers of the target language. To address, if not all, these weaknesses, I included two additional Google search engines that are more specialized: Google Scholar and Custom Search Engine (CSE). First, as Google Scholar searches only a database of scholarly literature, it can be used when the user wants to confine her queries to academic registers, which are assumed to contain much fewer lexical and grammatical errors than language data extracted from the entire Web may have. Second, Custom Search Engine, another search service provided by Google, searches only websites or URLs sharing specific patterns designated by the user. It can be used when the user wants to further narrow her searches down to a particular domain or academic discipline to get results that are more relevant and specific to the user’s own field of study. As seen in Figure 3.4, the Google tab has different search engine options other than Google (Web) and Google Scholar. The remaining engines are all CSEs customized for the present research participants. In the initial survey, the participants provided the names of major academic journals in their disciplines. Based on these, I created five CSEs named with an initial representing each participant’s field of study. Table 3.3 lists journals that were accessed for searches in each CSE featured in the Google tab.

Table 3.3

<table>
<thead>
<tr>
<th>Name</th>
<th>Field of study</th>
<th>Sites searched in each engine</th>
</tr>
</thead>
</table>

2 Two of the participants were in similar disciplines and therefore their CSEs were integrated into one. Hence, there are five CSEs.
Just The Word

The third tab in i-Conc is dedicated to another online concordancing web service, JustTheWord (JTW). The service (provided by Sharp European Research Lab) makes use of an 80 million-word subsection of the BNC. Using its own statistical methods and analysis technology—the website does not reveal exactly what methods they use to extract collocates—this resource provides collocational information, i.e., the company a word keeps in its vicinity and the strength of the word combinations they form. Unlike COCA, JTW has a relatively simple interface with only two major search options and requires no special query syntax.

![JTW combinations](image)

**Figure 3.5.** Screenshot of the results of the query “assumption” run in JTW combinations.

To briefly look at the two search options, first, a single-word query entered in *JTW combinations* typically returns its collocates extracted from the corpus, which are displayed in clusters grouped by the collocates’ parts of speech and senses. The strength of a combination that the queried word and its collocate form is represented by a green bar displayed next to each
collocation entry—i.e., the longer the bar is, the stronger the collocational relation is (see Figure 3.5).

The second search option *JTW alternatives* takes two to three-word queries (as opposed to single-word queries) and returns an indication of how strong the collocational relation between these query words is, which is again represented by the green bar, and provides other possible alternatives that can replace each constituent word in the queried combination (see Figure 3.6).

![Figure 3.6. Screenshot of the results of the query “make assumption” run in JTW alternatives.](image)

In JTW, as can be seen in Figures 3.5 and 3.6, word combinations displayed in the results page are not in their full forms used in actual sentences. They show only the main words (in their base forms, e.g., *make assumption*) that form the combination with no grammatical inflection or intervening words such as articles, determiners, or modifiers. A click on a word combination entry returns actual concordances in which the combination occurs with all the grammatical inflections and intervening words (see Figure 3.7).

![Figure 3.7. Screenshot of concordances of the word combination “make assumption”.](image)
As described above, JTW is a relatively simple tool useful especially when quickly checking collocates of a word and the strength of their combination.

3.2.4 Dictionaries

Dictionaries, the fourth and last tab in i-Conc, provides freely available online dictionaries that can complement the concordancers in the previous tabs. As can be seen in Figure 3.8, three dictionaries are accessed in this tab: Bilingual (Naver, an online bilingual Korean-English, English-Korean dictionary), LDOCE (an online version of the Longman Dictionary of Contemporary English), and Thesaurus (Roget’s 21st Century Thesaurus) (see Table 3.2 for URLs of these dictionaries).

![Figure 3.8](image.png)

Figure 3.8. Screenshot of the Dictionaries tab on i-Conc.

3.3 Procedures

As noted above, the present study was conducted in two phases. In the first phase of the study, the participants filled out an initial survey and had an interview about their approaches and attitudes toward English academic writing and their uses of reference resources. Then they received individual tutoring (three hours in total) from me, in which basic corpus linguistics concepts were introduced and hands-on practice in using each resource in i-Conc was provided. They had two to four weeks to use the tool independently whenever they needed to consult
reference resources. To examine whether the participants had gained enough familiarity with *i-Conc* as writing assistance, they were asked to choose a section from a paper they had written before and revise it using *i-Conc* for 50-60 minutes. While doing so, they recorded the screen activities using a screen recording program installed on their computers. I then conducted a stimulated recall with each participant, in which they shared the intentions behind each query and decision they made while watching the video together. At the end of the recall session, I provided feedback to each participant on how to improve his or her *i-Conc* consultation.

In the second phase of the study, the participants used *i-Conc* while completing an authentic writing assignment they chose and conducted three screen recordings (50 minutes or longer) of their writing process—twice at the drafting stage and once at the editing/revising stage\(^3\). Each screen recording was followed by a stimulated recall session as in the first phase. Finally, the participants had a final interview and filled out an evaluation survey, in which they shared what they perceived as the advantages and disadvantages of using the tool for their academic writing. Table 3.4 summarizes the procedures the participants followed.

Table 3.4

<table>
<thead>
<tr>
<th>Phase</th>
<th>Procedures</th>
</tr>
</thead>
</table>
| Phase 1 | - Initial survey and interview  
- 3 hours of training and hands-on practice  
- 2-4 weeks of independent use of *i-Conc*  
- Revision of a section from a course or research paper they previously wrote using *i-Conc* (50 minutes +, screen recorded)  
- Stimulated recall and feedback |
| Phase 2 | - Brief recap on *i-Conc* and orientation on Phase 2  
- Selection of an authentic writing assignment for the study  
- 3 screen recordings while completing the writing assignment using *i-Conc* (50 minutes + each)  
- Stimulated recall and feedback after each screen recording  
- Evaluation survey and interview |

3.3.1 Training

\(^3\) One of the assumptions made from the initial interviews was that referencing behaviors would be different depending on the stage of the writing process. To see possible differences caused by writing stages, the participants were asked to space their screen recordings this way.
The vast majority of previous studies pointed out how gradual and guided training would help learners exploit more fully corpus tools and other reference resources to support their task performance and learning. The importance of learner training in the use of new technology has long been recognized in the broader CALL context (Hubbard, 2004; Warschauer, 2005). Learners have been shown to be inconsistent in the use of CALL tools when left on their own and therefore they need guidance for the effective use of the tool and motivation to make needed investment in familiarizing themselves with it (Chapelle, 2003; Hubbard, 2013). As far as this study was concerned, this need for guidance was even more compelling as the participants were asked to independently use the reference suite that brings together multiple concordancing tools, with which most of them were not familiar.

I provided the participants with an average of three hours of individual tutoring, which was divided into two 1.5 hour sessions\(^4\). Three hours may not have been sufficient for them to reach an effective level of use for all the resources. But I considered three hours an appropriate amount of time for training with this specific group of participants for the following reasons. First, as graduate students, they had busy schedules with their own courses and research work. Some participants had a hard time finding even three hours for the training. Second, more importantly, care had to be taken not to overwhelm the participants with too much guidance on how and what to search and thereby to leave enough room for them to develop their own query strategies and skills to meet their needs.

The tutoring was broadly divided into three sections. First, the participants were introduced to basic concepts around concordancing. Secondly, they learned the characteristics of each resource and key query syntax and operators required to use in some of the resources. Finally, they were provided with ideas and tips on how to consult the suite in solving problems they encounter while writing in English. Table 3.5 provides a brief description of what was covered in the tutoring sessions. A more detailed description is also provided in Appendix B: \(i\)-Conc tutorial, which was provided for the participants at the beginning of the training as a manual to consult whenever they had conceptual and technical questions regarding \(i\)-Conc use.

\(^4\) Unlike other participants, Goeun received a single session of training, which was about two hour long, due to her tight schedule during her participation in the study. However, she was the only one among the participants who had prior experience of using corpus tools and already had considerable knowledge of corpus linguistic concepts and concordancing skills.
Table 3.5

*Content of Tutoring*

<table>
<thead>
<tr>
<th>Section</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: basic concepts</td>
<td>Introduction to corpus linguistics, its applications to language learning, learner concordancing; basic conceptual orientation on lexico-grammar: collocation, colligation, and semantic preference/prosody</td>
</tr>
<tr>
<td>2: <em>i</em>-Conc use</td>
<td>Basic purposes and functions of queries identified in previous studies; general instruction of <em>i</em>-Conc; the main characteristics of each resource; key query syntax and operators; hands-on practices with <em>i</em>-Conc</td>
</tr>
<tr>
<td>3: application to writing</td>
<td>Strategies for formulating queries to solve specific types of problems and for evaluating query results and applying them to their writing</td>
</tr>
</tbody>
</table>

As all the tutoring sessions were conducted individually, what was covered in the tutoring differed slightly depending on the participant. Each participant had different levels of linguistic and metalinguistic knowledge, so some parts of the training content described above were further elaborated on or shortened accordingly. It should be noted that the timing of study participation may have affected the effectiveness and quality of the training. The tutoring of the second group (Shia, and Ian), who received their tutoring in the summer of 2012, may have been qualitatively, if not quantitatively, better than that of the first group (Jae, Yumee, Jinho, and Goeun), who had their training in the winter of 2011-2012, for two reasons. When the second group joined the study, *i*-Conc had full functionalities with a few updates for its features. It was not the case with the first group, as the tool had not then been fully developed, and therefore their hands-on practice during Phase 1 was somewhat limited compared to the second group. I, as the researcher, also gained new insights into search strategies and effective uses of each resource after working with the first group. I integrated these new insights into the tutoring for the second group.

3.3.2 Data collection

Data collection was conducted using multiple sources and methods in a mixed methods design in order to answer the multi-faceted research questions that required eliciting and analyzing the cognitive processes the participants went through while interacting with *i*-Conc as
well as factors that might influence individual differences. In what follows, each data source is discussed in detail.

3.3.2.1 Surveys

At the beginning of the study, the participants were surveyed on their backgrounds (age, major, degree pursued, years in Canada, etc.), computer literacy, familiarity with corpus tools, attitudes and approaches to writing in English, difficulties they experience in English writing, and the reference resources used when writing and the purposes of their reference resource consultation. At the end of the study, another survey was conducted to elicit the participants’ evaluations of the tool as writing assistance in respect to: what they saw as the benefits and limitations of concordancing and dictionaries in different aspects of writing, the extent to which the tool use helped to increase their autonomy and confidence as L2 writers, whether they had changed their attitudes and approaches to English writing, and their willingness for future use of the tools. Both surveys were conducted online using the free service provided by eSurveyspro.com (see Appendix C for the questionnaire).

3.3.2.2 Interviews

The participants each had a 40 to 50 minute interview after completing the initial survey questionnaire. In this initial interview, the participants shared in detail their perceptions and processes of writing in English and attitudes toward it, areas of difficulties, and information about the sources and types of writing support and reference resources they usually consulted. At the end of each stimulated recall session, the participants also had informal interviews of varying lengths on their ongoing experiences of using i-Conc as writing assistance. At the end of the study, the participants also had in-depth interviews, for about 40 to 50 minutes, in which they provided their overall evaluations of the reference suite and described the changes in their attitudes and approaches to English academic writing. The interviews were semi-structured to create room for the participants to elaborate on what they thought was particularly important and to offer insight into aspects of the research that I may have neglected (the questions asked during the initial and final interviews appear in Appendix D). All the interviews were done in Korean and audio recorded using a digital recorder. I transcribed the initial and final interviews fully. For the short informal interviews done at the end of the recall sessions, I transcribed only
the parts that I judged were important. I also translated into English all excerpts from the participants’ interviews and surveys that appear in the thesis.

3.3.2.3 Writing assignments

The main focus of the study was to examine the independent uses of a suite of reference resources by Korean ESL students for supporting their authentic writing tasks for academic purposes in English. There were no specific writing tasks participants had to carry out solely for purposes of the study. Rather, participants were asked to record and report on their uses of i-Conc while completing their authentic writing tasks such as a term paper for one of the academic courses they were taking.

There were two writing samples collected from each participant. In the first phase of the study, after the training sessions and familiarization period, the participants screen-recorded for 50 minutes or longer their own process of revising or editing a portion of the paper they had written before. These papers were collected for analysis. In the second phase, the participants selected a major writing assignment (at least 10 pages long) such as a course paper or research proposal or presentation, and used i-Conc over the course of completing the writing assignment from drafting to editing and revising. These papers were also collected for analysis. Table 3.6 provides descriptions of the types of writing assignments the participants selected for the present study.

Table 3.6

Writing Assignments

<table>
<thead>
<tr>
<th>Participant</th>
<th>Writing Assignment (length limit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jae</td>
<td>Conference paper (30 pages)</td>
</tr>
<tr>
<td>Yumee</td>
<td>Introduction and literature review sections of her dissertation proposal (20)</td>
</tr>
<tr>
<td>Jinho</td>
<td>Term paper for one of his courses (10)</td>
</tr>
<tr>
<td>Goeun</td>
<td>Term paper for one of her courses (15)</td>
</tr>
<tr>
<td>Shia</td>
<td>Term paper for one of her courses (10)</td>
</tr>
<tr>
<td>Ian</td>
<td>Dissertation proposal (40)</td>
</tr>
</tbody>
</table>
As seen in Table 3.6 above, three participants (Jinho, Goeun, and Shia) chose a course paper for the study with a tight deadline compared to the other three participants (Jae, Yumee, and Ian), whose writing assignments were either a conference paper or dissertation proposal, and thus longer in length and had more flexible deadlines.

3.3.2.4 Query log

To trace participants’ look-ups and the processes in which they interacted with the reference suite, i-Conc was designed to send every query participants perform in its search box to a server and save it in a query log, which was developed by the same programmer who developed i-Conc. For each query, the query log displays the following information: (a) user ID, (b) time, (c) query, and (d) resource consulted. This information could be retrieved in its entirety or by individual categories (see Figure 3.9).

![Figure 3.9. Screenshot of query log.](image)

3.3.2.5 Screen recording and stimulated recall

Following the methodology used in Park (2010), I used both screen recording and stimulated recalls as methods of participant observation. Screen recording refers to real-time recording of the computer screen activities, which captures everything that is displayed on the screen as a video file. Screen recording software runs in the background on the user’s computer, so this method allowed me to conduct unobtrusive observation of participants’ behavior as has often been adopted in other recent concordancing studies (Hafner & Candlin, 2007; Kennedy & Miceli, 2010; Lee, Lin, & Liou, 2006; Park, 2010).

For the present study, the participants installed on their own laptop computers a free screen recorder, BB Flashback Express.
did the recordings at times of their choosing, and uploaded the video files onto their individual folders created in my personal Web storage site, which only I could access. Figure 3.10 is a screenshot of a screen-recorded video of a participant’s writing.

I watched each video and compared it with the query log and created a full list of queries performed during the recording. Each participant and I conducted a stimulated recall watching the video together within 1 to 2 days of the recording. Participants provided their recollections of why they performed each query on the list, explaining how they decided whether to apply search results to their writing, thereby providing clues to understanding the cognitive processes behind their interactions with the tool, and more specifically to determine their problem solving motivation, the types and purposes of queries made, and strategies used. The guide for these stimulated recall protocols appears in Appendix E.

Each participant carried out four sets of a screen recording and a follow-up stimulated recall based on it. Among these four sets, the first was conducted during the revision task in the first phase (SR1), the second and third at the drafting stage (SR2, SR3), and the last one at the editing/revising stage (SR4) of the main task in the second phase. The participants produced a

---

Most participants reported in the initial interviews that their writing process was not linear and sequential going through well-defined stages but simultaneous and recursive as is well documented in writing research (Flower & Hayes, 1981). Participants were thus asked to space their screen recordings so that the first two recordings would be done while drafting is a dominant activity and the last one near the completion of the paper where they engage mainly in editing and revising.
total of 1,509 minutes of video data through 24 screen recordings (4 videos x 6 participants). Table 3.7 provides the length of each screen recording.

Table 3.7

<table>
<thead>
<tr>
<th>Participant</th>
<th>SR1</th>
<th>SR2</th>
<th>SR3</th>
<th>SR4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jae</td>
<td>51</td>
<td>58</td>
<td>56</td>
<td>50</td>
<td>214</td>
</tr>
<tr>
<td>Yumee</td>
<td>67</td>
<td>68</td>
<td>82</td>
<td>49</td>
<td>263</td>
</tr>
<tr>
<td>Jinho</td>
<td>27</td>
<td>60</td>
<td>57</td>
<td>58</td>
<td>202</td>
</tr>
<tr>
<td>Goeun</td>
<td>66</td>
<td>64</td>
<td>61</td>
<td>39</td>
<td>230</td>
</tr>
<tr>
<td>Shia</td>
<td>37</td>
<td>51</td>
<td>67</td>
<td>43</td>
<td>198</td>
</tr>
<tr>
<td>Ian</td>
<td>61</td>
<td>164</td>
<td>81</td>
<td>97</td>
<td>402^a</td>
</tr>
</tbody>
</table>

Note. a. Ian’s screen recordings were much longer than other participants’. Instead of stopping recording after the required time length, which most other participants did, he recorded throughout his writing sessions from beginning to end.

3.4 Data analysis

To answer the research questions, which were both quantitative and qualitative in nature, multiple analyses were conducted. To give an overview of the analyses done for the study, firstly, units of analysis were identified to segment the participants’ i-Conc use data. Coding schemes were then developed to classify each participant’s data (screen recordings, stimulated recall protocols, and query log data) into the units of analysis identified and measure the participants’ problem solving performance. Secondly, interviews, surveys, and writing samples were analyzed qualitatively to examine the participants’ evaluations of i-Conc as writing assistance and its impact on their processes, approaches, and confidence regarding their English academic writing. Especially relevant themes and factors that potentially represented individual differences were identified to carry out cross-case analysis. In what follows, the data analysis procedures are discussed in detail in the order described above.

3.4.1 Units of analysis

As in most previous research on reference resource consultation (Kennedy & Miceli, 2001; Park, 2010; H. Yoon, 2005) and more broadly text formulation while writing (Cumming,
I considered i-Conc consultation the problem solving activities the L2 writer engages in to search for solutions to a wide range of problems encountered while writing in the target language.

Because a main focus of the study was to examine the types of problems participants take up to solve using reference resources, not all writing problems they experienced while completing their assignment were identified or analyzed. Instead, only linguistic writing problems that prompted i-Conc consultation—identified through triangulating data from the screen recordings, the query log, and the participants’ recollections—were examined. This criterion of triangulation meant that the queries that were recorded in the query log but not screen-recorded were excluded from the analysis.

A preliminary analysis showed that a linguistic writing problem can be resolved (or abandoned) with a single look-up in a reference resource but in general requires multiple look-ups in the tool to arrive at a final decision on a given linguistic item searched. To examine sub-units within a problem, I used the concept of problem space (Newell & Simon, 1972). In cognitive psychology, a problem space is defined as all possible states and operations the problem solver goes through to reach a solution to the problem at hand (Anderson, 1995; Heine, 2010). Put differently, a problem space can be visualized as a space consisting of an initial state, which the problem solver finds unsatisfactory, and a desired goal state with intermediate states and steps located between the two points. Applying this concept, all the single look-ups carried out in the process of solving a specific problem can be seen as problem states or operations leading up to the desired goal state, which as a whole constitutes a problem space. Thus, the unit of analysis for i-Conc consultation data is the problem space and its subunits are single look-ups making up the space. In this study these single look-ups are referred to as queries, and a problem space is defined as all the queries the participant performs for the resolution of a problem using i-Conc. An example of a problem space as the unit of analysis with all its sub-units (i.e., queries) is provided in Appendix F.

3.4.2 Coding scheme – types of problems and queries

To get answers to the first research question, which is concerned with the purposes and types of i-Conc consultation, I developed a coding scheme to break the units (i.e., problems and
queries) identified further down according to their types. The scheme was built from consideration of previous studies (Frankenberg-Garcia, 2005; Kennedy & Miceli, 2010; Murphy & Roca de Larios, 2010; Park, 2010; Roca de Larios et al., 1996) and also from the preliminary analysis of my actual data.

At the broadest level, problems are classified into two categories according to their initial problem solving motivation. First, a confirmatory problem refers to instances when the writer has a specific linguistic item\(^6\) in mind or in writing for a meaning she intends to express and tries to confirm whether they match accurately and/or appropriately in a given context. A compensatory problem, on the other hand, refers to instances when the writer does not have a clear idea about how to convert into text a meaning she intends to express and tries to extract potential target items from the reference resources.

At the next level, queries within a problem space are classified into two types: verification and elicitation. A verification query is one performed to find answers to a closed-form question the writer poses (i.e., yes/no questions) (Kennedy & Miceli, 2001) and to verify—hence the name—from the query results the presence or frequency of the instances of an item or its lexical, grammatical, and/or stylistic accuracy and appropriacy in a given context—e.g., is the adverb deeply the most typical collocate of the word entrenched? An elicitation query is, in turn, one performed to find answers to an open-form question (i.e., a what or how question) and usually provides the problem-solver with more than one choice in the query results—e.g., what are the typical adverbs that modify the word entrenched? Therefore, it is used to elicit from the resources possible candidates for the target item. Problem and query types are closely related. In general, a compensatory problem solving cycle is usually started with an elicitation query to extract from i-Conc what is unknown to the problem-solver whereas a confirmatory cycle is initiated with a verification query to test a hypothesis made based on the problem-solver’s existing knowledge. However, a problem space initiated by one type of motivation, either confirmatory or compensatory, can contain both verification and elicitation queries. For example, in a typical confirmatory problem space, if the initial verification query returns a negative answer—i.e., the results confirm that the item in mind does not convey the writer’s intended meaning—the writer tends to perform an elicitation query to retrieve alternatives.

---

\(^6\) Items here refer to linguistic elements at various levels ranging from morphosyntactic, lexical, syntactic to discoursal or stylistic elements.
At another level, queries are also classified according to their specific purposes. Starting from the categories used in previous studies, query purposes were identified through several iterations of analyzing screen recording and stimulated recall data. Although the participants used i-Conc for many different purposes, I decided to include 14 purposes for coding, each of which occurred more than three times across three participants and more. The purposes that did not satisfy this condition were classified as others. Whereas the distinction between verification and elicitation queries lies in the formal aspect of the question that prompts the query (whether vs. what or how), the purpose of a query is determined by the content of the question. The 14 query purposes are provided along with typical questions that prompt a query for each purpose in Table 3.8. Finally, each query was classified by the linguistic nature of the queried item: lexical, grammatical, lexico-grammatical, or stylistic. Definitions of these categories are also provided in Table 3.8. It should be noted that at the query levels multiple codes can be assigned to a single query as a query may have multiple purposes and belong to more than one linguistic category.

Table 3.8

**Coding Scheme for Problems and Queries**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Problem – initial motivation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confirmatory</td>
<td>The writer has a specific linguistic item in mind or in writing for a meaning she intends to express and tries to confirm whether they match accurately and/or appropriately in a given context.</td>
<td>Conf</td>
</tr>
<tr>
<td>Compensatory</td>
<td>The writer does not have a clear idea about how to convert her intended meaning into text and tries to extract potential target items from the reference resources.</td>
<td>Comp</td>
</tr>
<tr>
<td><strong>Query – type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verification</td>
<td>Performed to verify the presence/frequency of the instances of an item or its lexical, grammatical accuracy, and/or stylistic appropriacy in a given context. Typically prompted by a “whether” question the writer poses – e.g., is the adverb “deeply” is the most typical collocate of the word “entrenched”?</td>
<td>V</td>
</tr>
<tr>
<td>Elicitation</td>
<td>Performed to elicit linguistic items that match the writer’s intended meaning and typically prompted by a “what” or “how” question – e.g., what are the typical adverbs that modify the word “entrenched”?</td>
<td>E</td>
</tr>
<tr>
<td><strong>Query – purpose</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple confirmation</td>
<td>Is X (a word, phrase, or syntactic pattern) actually used? Can I say X in this context? Is X typically and frequently used in the given context?</td>
<td>SC</td>
</tr>
<tr>
<td>L2 equivalent</td>
<td>What is the L2 equivalent of X?</td>
<td>Equi</td>
</tr>
<tr>
<td>Collocation</td>
<td>Is X a typical or appropriate collocate of Y? What is the most typical or appropriate collocate of X?</td>
<td>Coll</td>
</tr>
<tr>
<td>Intended meaning</td>
<td>Does X deliver the intended meaning? What is the accurate meaning of X in the given context?</td>
<td>Mn</td>
</tr>
<tr>
<td>Simple alternatives</td>
<td>What are the alternatives that can replace X to avoid repetition?</td>
<td>sAlt</td>
</tr>
<tr>
<td>Upgrading alternatives</td>
<td>What are the alternatives of X that are more accurate and appropriate in the given context? How can X be better expressed?</td>
<td>uAlt</td>
</tr>
<tr>
<td>Article/determiner</td>
<td>Should X be used with an article/determiner?</td>
<td>Art</td>
</tr>
<tr>
<td>Argument pattern</td>
<td>What are the typical/frequent object types X takes? Does X take a noun (phrase), to infinitive, gerund or clause as its object? Is X a transitive or intransitive verb?</td>
<td>AP</td>
</tr>
<tr>
<td>Noun ending</td>
<td>Should X be in the singular or plural? Is X a countable or uncountable noun?</td>
<td>NE</td>
</tr>
<tr>
<td>Sentence/phrases hunting</td>
<td>How can multiple content words be combined in a sentence? (searching for sentences similar to what the writer is trying to say)</td>
<td>SH</td>
</tr>
<tr>
<td>General usage</td>
<td>How is X used in sentences? (aimed to discover typical syntactic patterns in which it is used or its collocation)</td>
<td>GU</td>
</tr>
<tr>
<td>Register</td>
<td>Is X in the right register? (formal/informal, or academic/non-academic)</td>
<td>Reg</td>
</tr>
<tr>
<td>Part of speech</td>
<td>Is X a noun (or a verb, etc.)? Which word class does X belong to?</td>
<td>POS</td>
</tr>
<tr>
<td>Spelling</td>
<td>Is X spelled right? How is X spelled?</td>
<td>Sp</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td>Oths</td>
</tr>
</tbody>
</table>

**Query – linguistic category**

**Lexical**
Queries regarding the meaning, presence, frequency/typicality, accuracy/appropriacy, etc. of a lexical item—a single word or phrase that can be treated as a lexical unit (Lewis, 1997)—in a given context

**Grammatical**
Queries regarding syntax (structure), morphological problems (word form), tense, number, article/determiner, etc.

**Lexico-grammatical**
Queries regarding collocation, and colligation between and across words and grammatical structures, in which grammar and lexis constrain each other; queries about general usage of a word or phrase

**Stylistic**
Queries regarding register, formality, rhetorical effects, etc.

**Note.** a. Collocations here include both lexical (i.e., combinations among nouns, verbs, adjectives, and adverbs) and grammatical (i.e., verbs, adjectives or nouns combined with a preposition or a grammatical structure) collocations, which roughly follows the categorization of Benson, Benson, and Ilson (1986).

3.4.3 Coding scheme – problem solving results
To find answers to the second research question regarding the effects of i-Conc use on the participants’ problem solving performance, I developed another coding scheme. Specifically, to examine the extent to which the participants were satisfied with i-Conc consultations they performed and how these affected their actual problem solving performance, each problem space identified above was coded on the following three dimensions: (a) whether the writer was satisfied with the consultation results—i.e., whether she perceived as satisfactory the answer she found to the problem that had prompted the consultation, (b) whether the writer actually found a correct solution to the given problem—i.e., a successful verification of the syntactic, semantic or stylistic correctness of the given item or a correct text formulation/revision of the given ideational content—and (c) whether the writer abandoned the search and formulated text with no input from i-Conc or abandoned both the search and text formulation of the given ideational content altogether. Coding data on these three dimensions adapted and expanded upon the procedures in Frankenberg-García’s (2005) reference resource consultation study. Codes on the first dimension are intended to show the participants’ perceived helpfulness of the tool in resolving each problem they tackled. Based on the participants’ recall protocols, each problem space was rated as “Satisfied” (S) or “Dissatisfied” (D). The second dimension reveals whether the participant’s problem solving efforts actually led to the resolution of the problem. The solutions the participant applied to the problems were evaluated as “Correct” (C) or “Incorrect” (I). Instances where the participant gave up on the text formulation of a given ideational content altogether were coded as “Not Applicable (N/A)”. Finally, the third dimension categorizes whether, faced with unsatisfactory consultation results, the participants gave up performing further queries without retrieving anything from the consultation results. These problem spaces discarded midway were coded as “Abandoned” (A). Coding on these three dimensions is intended to illustrate how successfully the participants carried out i-Conc consultations, how it compared with the participants’ perceptions of their own consultation results, and how they dealt with the problems they gave up solving with i-Conc. Table 3.9 presents descriptions and definitions of these categories and codes.
Table 3.9

*Coding Scheme for Problem Solving Results*

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Satisfaction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfied</td>
<td>For a given problem, the participant is satisfied with the solution she finds from ( i)-Conc search results and applies it to his or her writing.</td>
<td>S</td>
</tr>
<tr>
<td>Dissatisfied</td>
<td>For a given problem, the participant is not satisfied with the search results as the consultation does not return a clear or immediate solution to the problem. There are two paths the participant can take. First, the participant retrieves what she thinks may be a solution to the problem and applies to the writing, but she is not sufficiently confident in its accuracy or appropriacy. Second, the participant gives up performing further queries and retrieves nothing from the search results.</td>
<td>D</td>
</tr>
<tr>
<td><strong>Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>For a given problem, the participant finds a correct and appropriate solution and applies it to the writing.</td>
<td>C</td>
</tr>
<tr>
<td>Incorrect</td>
<td>The solution the participant applies to a given problem is semantically, or syntactically incorrect, or stylistically inappropriate in the given context.</td>
<td>I</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>The participant gives up formulating the given ideational content into text.</td>
<td>N/A</td>
</tr>
<tr>
<td>Abandoned</td>
<td>The participant is dissatisfied with the search results at some point within a problem space and gives up performing further queries.</td>
<td>A</td>
</tr>
</tbody>
</table>

### 3.4.4 Case studies and cross-case analysis

To address the research questions of a more qualitative nature (research question 3 and 4), each participant’s data were analyzed for the following as an individual case study:

a. a general problem solving process using \( i\)-Conc, or \( i\)-Conc consultation sequences,

b. individual participants’ specific \( i\)-Conc use patterns—strategy uses, challenges and difficulties encountered, etc.—at different stages of formulating queries, interpreting/evaluating the search results and applying the results to their writing,

c. perceived changes in the processes and approaches regarding their English academic writing while using \( i\)-Conc, and

d. the participants’ evaluation of \( i\)-Conc (the tool and individual resources featured in it) as a writing supporting tool and the impact of its use on their confidence in English academic writing.
For analyses a and b, screen recording, stimulated recall, query log, and writing sample data were triangulated. For analyses c and d, surveys, initial and final interviews, and informal interviews conducted at the end of stimulated recalls were coded through several iterations of thematic coding.

Finally, to put findings from each analysis in perspective and thus to gain a comprehensive picture of what influenced individuals’ interactions with i-Conc and evaluation of it and thereby to answer the fifth research question regarding possible factors affecting individual differences, I conducted a cross-case analysis synthesizing the individual case studies. Specifically, I consolidated the common and salient themes that emerged from the participants’ survey questionnaire and interview data into eight broad factors corresponding to the three fundamental dimensions of writing: writer, text, and context (cf. Cumming, 1990, 1998). Individual participants (cases) were compared and contrasted to see how these factors interacted with one another and how differences in these factors played out in the participants’ actual i-Conc use patterns and evaluation.

3.5 Reliability/credibility

To ensure the overall reliability and credibility of findings and inferences based on them, the study employed a mixed methods design, in which data were collected from multiple sources and analyses were conducted using multiple methods as described above. There are many purposes for using a mixed methods design but one overarching rationale is that when studying a complex phenomenon, a mix of methods will yield better and broader understandings than a single method alone (Dörnyei, 2007; Greene, 2006). Given the complex and multi-faceted nature of reference resource consultation while writing and issues around it, multiple methods in the present study were pursued on the following two dimensions: for the participants’ i-Conc use itself, the study attempted to get converging and corroborated findings through data triangulation, observing participants’ interactions with the tool from multiple angles such as screen recordings, a query log, and participants’ recollections of their tool use. For the issues surrounding i-Conc use, meanwhile, the study sought a comprehensive and deep understanding through a complementary mix of different methods such as surveys, interviews, observation,
and analysis of writing samples. The findings from these mixed methods analyses are in turn presented in three different and complementary ways in the chapters to follow: group findings (Chapter 4), individual cases (Chapter 5), and intervening factors (Chapter 6).

To establish reliability of data coding in particular, two additional coders, both PhD students in second language education, participated in the coding. Each of the problems and queries identified was assigned a unique ID number on an Excel spreadsheet, and 20% of them were randomly selected using Excel formulas (RAND and INDEX). After having an individual practice session, one of the coders coded independently this randomly selected 20% according to the coding scheme presented in Table 3.8. The other coder likewise coded about 20% of the problems for problem solving results following the coding scheme shown in Table 3.9. The inter-coder agreements were high: 100% for problem types and 93%, 91%, and 84% for query types, purposes, and problem solving results respectively.

Another step taken to ensure the credibility of data analysis was member checking. Once I drafted case studies, I sent them out to the participants via email asking each participant to read his or her case study report and assess the accuracy and adequacy of the inferences made in the report and to elaborate and provide additional information if he or she found the report lacking and inadequate. Three out of the six participants responded with feedback, which mostly concerned minor inaccuracies in their background information reported in the case studies. The case studies were revised reflecting their input.

3.6 Notation convention

In the chapters to follow, there will be a great number of exemplifications and categorizations referring to particular problem spaces and queries within them. I therefore developed a notation convention to cite the problem spaces and queries. Each problem space identified has been assigned a unique ID (e.g., YP05), which consists of two parts (YP + 05): The first part consists of a letter from the participant’s name and the letter “P” representing the word “problem”, while the second part represents a serial number given to each in order of its appearance in the participant’s screen recordings (SR1-4). YP05 here thus refers to the fifth problem in Yumee’s data. The first parts of the IDs for the other participants are as follows: JP (Jae), GP (Goeun), HP (Jinho), SP (Shia), and IP (Ian). A particular query within a problem
space in turn is expressed with a number assigned in order of its appearance in the given problem space. This number is attached to the ID of the problem space. YP05-2, for example, refers to the second query of problem space YP05.
CHAPTER FOUR

GROUP FINDINGS

In this chapter, findings from the analyses described in the previous chapter are presented. The chapter is structured into five sections. The first section presents findings from the initial survey and interviews to contextualize the quantitative and qualitative findings that follow, while the next four sections address findings related to each research question posed in Chapter 1.

4.1 Profiles of the participants

4.1.1 Experience and perceptions of English academic writing

Being graduate students in writing-intensive disciplines at a North American university, most of the participants had already had considerable experience writing academic papers. However, there was also great variability in the amount and type of academic writing they had done. Although all four PhD students wrote their MA theses in English back in Korea, some had greater experience than others as they were at the different stages of their PhD studies. For example, Jae, who already finished his course work, had written dozens of course papers and conference papers, while Goeun, who had just started her PhD program, had written only one paper in the previous semester. The two MA students were also at different stages. Shia was in her last semester and had written many course papers in her three previous ones.

Table 4.1

Participants’ Prior Experience of Academic Writing

<table>
<thead>
<tr>
<th>Jae</th>
<th>Yumee</th>
<th>Goeun</th>
<th>Jinho</th>
<th>Shia</th>
<th>Ian</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA thesis, dozens of major course papers, several conference papers</td>
<td>MA thesis, 12 major course papers, 5 conference papers</td>
<td>MA thesis, 2 major course papers</td>
<td>Course papers throughout undergraduate years, 4 course papers in his MA program</td>
<td>Dozens of major course papers in her MA program</td>
<td>MA thesis, 8 major course papers</td>
</tr>
</tbody>
</table>
Although Jinho had earned his undergraduate degree in Canada and therefore written quite a lot of course papers, he, then in the second semester of his current program, had been struggling with different levels of expectations for his academic papers. Table 4.1 presents the types and amount of academic writing in English the participants had completed before taking part in the study.

 Asked about their attitudes toward English academic writing, three participants (Jae, Goeun, and Ian) responded that they were often frustrated by its time-consuming nature and their inability to express what they intend to say as accurately and appropriately as they would do in their L1. By contrast, Yumee answered that she was enjoying academic writing while Shia and Jinho responded they had no strong attitudes, neither positive nor negative, toward English academic writing. As to confidence in writing academic papers in English, most participants rated their confidence as low because of their perceived lack of sufficient proficiency in English. However, Yumee responded that although she was not highly confident about lexical or grammatical accuracy, she did not find English academic writing particularly challenging or difficult as all her exposure to the related content was in English. Shia also said that she did not get nervous or insecure about her English writing anymore because she believed her papers were good enough in terms of organization and coherence although lacking in vocabulary and grammar.

 As for specific aspects of writing that they found challenging, all participants cited repeated use of a limited repertoire of vocabulary and syntactic structures as their main areas of difficulties in English academic writing. Other areas of difficulties included expressing subtle and fine nuances (Goeun, and Ian), academic style or register (Jae), organization and coherence (Jinho), and specific grammar items such as articles and prepositions (Shia). In the survey, they rated on a scale of 1 (not difficult at all) to 5 (very difficult) different aspects of academic writing, ranking Vocabulary/Usage as the most difficult ($M = 3.83$), followed by Grammar/Sentence Construction ($M = 3.33$), Rhetoric/Organization ($M = 3.33$), and Register/Academic Style ($M = 3.17$). Asked about which of the four aspects of English academic writing they hope to improve the most, half the participants chose Vocabulary/Usage, with the remaining three aspects each chosen by one participant respectively.

 As for specific aspects of writing that they found challenging, all participants cited repeated use of a limited repertoire of vocabulary and syntactic structures as their main areas of difficulties in English academic writing. Other areas of difficulties included expressing subtle and fine nuances (Goeun, and Ian), academic style or register (Jae), organization and coherence (Jinho), and specific grammar items such as articles and prepositions (Shia). In the survey, they rated on a scale of 1 (not difficult at all) to 5 (very difficult) different aspects of academic writing, ranking Vocabulary/Usage as the most difficult ($M = 3.83$), followed by Grammar/Sentence Construction ($M = 3.33$), Rhetoric/Organization ($M = 3.33$), and Register/Academic Style ($M = 3.17$). Asked about which of the four aspects of English academic writing they hope to improve the most, half the participants chose Vocabulary/Usage, with the remaining three aspects each chosen by one participant respectively.

 Asked whether they usually set a goal for an academic paper that they have to write and if so, what overall and specific goals they have, most participants responded that they rarely set
specific goals for different academic papers. For most of them, meeting the deadlines was the first and foremost goal as they usually had to deal with multiple papers at a given time in an academic term. Of particular note is that all of the participants said they never made it a goal to produce a grammatically or lexically error-free paper, which they saw as impossible for L2 writers. Rather, they focused more on organization (e.g., a thesis sentence + supporting statements), logical flow, and engaging content. Jae and Yumee in particular shared the reason why they try not to expend too much effort on getting the appropriate grammar and vocabulary while writing English papers: They both said that they always considered their papers “drafts”, which would often be merged with other papers and expanded into conference or journal papers through multiple revisions. Always pressed for time, they not only could not afford to spend much time and energy on a particular paper, but they saw it inefficient to do so for what would eventually be revised later.

4.1.2 Writing support and reference resource use

Asked what kind of support they usually receive for their English academic writing, the participants responded that their sources of writing support were writing centres, classmates, and professors. Five of the six participants said they had tried to take every academic paper they wrote to the writing centres to get some feedback on it. The only exception was Goeun, who had a tendency to complete her writing assignments right before the deadlines, and therefore had no time to get any feedback from peers or writing centre tutors. Those who got support from the writing centres commonly found the writing support they received insufficient but for different reasons. It was especially not possible for Shia and Jinho to get sufficient help for every paper they wrote because of the tight restrictions imposed by their faculty’s writing centres on the frequency and duration of tutoring service. Their tutoring service was typically restricted to once a week, and 45 to 50 minutes a session, and coupled with their own time availability, one session was usually the maximum they could allot to a paper before its deadline, which was often too short to discuss the whole paper. For Ian, what was lacking was the quality, rather than the amount, of feedback he got at the writing centre. He pointed out that the feedback the tutors provided was often “template like,” nothing beyond the basic rules and principles one would typically find in any academic writing textbooks, without much practical help with surface-level
The interviews also revealed that some participants used writing centre services strategically, making the most out of them under the given restrictions. Through years of using the writing centre services, Jae and Shia developed close acquaintances with a few specific tutors. They would book tutoring sessions only with these tutors, who, knowing their personal writing styles, and their weaknesses, provided feedback effectively and efficiently in the short time allowed. On their part, Jae and Shia also brought their papers with specific questions to ask by underlining parts they wanted the tutors to focus on.

The participants indicated they received feedback on their papers from their professors usually in the form of evaluation. The type and amount of the feedback from professors varied but it tended to focus more on overall content rather than linguistic aspects of writing, which is nothing unusual in North American universities. The only exception was Yumee. She said she was so lucky to have a supervisor who was willing to read every major paper she wrote, even ones written for the courses with other professors, and provide feedback with detailed suggestions about surface level editing. Yumee was also the only participant who would finish drafts of her papers well before the deadlines and ask her peers to review them and provide feedback.

Writing support from writing centres, professors, and classmates were provided most of the time towards the end of or after completing a paper. When immediate help was needed for specific linguistic (lexical and grammatical) problems frequently encountered while writing, all participants said they consulted dictionaries and other reference resources. The most commonly used reference resources among the participants were Korean-English bilingual, English-Korean bilingual, and English monolingual dictionaries. Google search engine and thesauri were used by three participants respectively. In addition to these resources, Ian and Goeun also used a collocation dictionary and Jinho said he consulted grammar books while writing. Goeun claimed to be previously using the greatest range of reference resources including online corpus tools, some of which feature in i-Conc. Other participants had no prior experience of using corpus tools or concordancers except for Google. Table 4.2 lists each participants’ sources of writing support and reference resources they said they use for writing.

The range of purposes they used these resources for varied depending on the participant but commonly included checking meanings, examples, and grammar related to a word or phrase,
finding an English equivalent to a Korean word or phrase, and confirming whether a linguistic item is used by native speakers. Most participants had mixed feelings toward the resources they were using. That is, while they were not fully confident about the correctness of answers they retrieved from them, they resorted to them as they were often the only resources they could turn to while writing with no other immediate writing support available.

Table 4.2

*Sources of Writing Support and Reference Resources*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Sources of writing support in linguistic aspects</th>
<th>Reference resources used while writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jae</td>
<td>Writing centre</td>
<td>K(^a)-E(^b) dict., E-K dict., E-E dict., Thesaurus</td>
</tr>
<tr>
<td>Yumeee</td>
<td>Writing centre, classmates, supervisor</td>
<td>K-E dict., E-K dict., Google</td>
</tr>
<tr>
<td>Goeun</td>
<td>Friends in the US only when desperate</td>
<td>K-E dict., E-K dict., Google, Collocation dict., Corpus tool, Thesaurus</td>
</tr>
<tr>
<td>Jinho</td>
<td>Writing centre</td>
<td>K-E dict., E-K dict., E-E dict., Google, Grammar books</td>
</tr>
<tr>
<td>Shia</td>
<td>Writing centre, a friend in the US</td>
<td>K-E dict., E-K dict., E-E dict., Google</td>
</tr>
<tr>
<td>Ian</td>
<td>Writing centre, near-native speaker friends</td>
<td>K-E dict., E-K dict., E-E dict., Collocation dict., Thesaurus</td>
</tr>
</tbody>
</table>

*Note.* a. Korean, b. English

In summary, most participants, despite their considerable writing experience and relatively high proficiency in English, still struggled with vocabulary and grammar while writing in English, notably with their perceived inability to match their intended meanings with the right lexical, grammatical, and stylistic forms in English. As a group, they considered Vocabulary/Usage the area that needed to be improved the most. To some degree, they tried to solve problems of these types that arose while writing by consulting linguistic writing resources but most of them believed that being non-native speakers, no matter how much time and energy they spent on vocabulary and grammar, they would not be able to get it “perfectly right”. They chose to focus more on non-linguistic aspects of their writing and instead sought help from native speakers such as writing centres and student peers, even though these human resources were not as available or satisfactory as they might have wished.

4.2 How and for what participants consulted *i*-Conc
This section presents answers to Research Question 1 by (a) describing the typical sequence of i-Conc consultation and (b) reporting descriptive statistics on problems and queries identified. From the section onward, many problem spaces and queries are cited as exemplification. See 3.6 Notation Convention for how specific problems and queries are represented.

4.2.1 Typical sequence of i-Conc consultation

In considering reference resource consultation a problem solving activity, I have correspondingly described the process or sequence of i-Conc consultation from the perspective of the problem solving cycle established in cognitive psychology. Although there are minor differences in the labeling of constituent steps, the problem solving cycles that commonly appear in cognitive psychology textbooks (e.g., Pretz, Naples, & Sternberg, 2003; Sternberg, Sternberg & Mio, 2012) have in common the following steps (adapted from Sternberg & Sternberg, 2012, pp. 445-446):

1. Problem identification: Do I actually have a problem?
2. Problem definition and representation: What exactly is the problem?
3. Strategy formulation: How can I solve the problem?
4. Organization of available information: How do pieces of information about the problem fit together?
5. Allocation of resources: How much time, effort, etc. should I put into this problem?
6. Monitoring: Am I on the right track in solving the problem?
7. Evaluation: Did I solve the problem correctly?

This cycle is a generalized model that can apply to a wide range of problem types and problem solving situations, though some of the steps in this cycle may be reversed in order, combined, skipped or repeated depending on the specific problem and problem solver (Sternberg et al., 2012). Particularly when the problem to solve is ill-defined, or ill-structured, the problem may have no clear solution paths or a possibility of multiple solutions or no solution at all (Jonassen, 2003), so the problem solver may have to reiterate the intermediate steps from 2 to 6 (or some of them) until she finds a satisfactory solution or otherwise gives up the problem solving.
When applied to reference resource consultation, the cycle can have fewer steps as the range of problems tackled are relatively narrow and concrete, but some of them are often reiterated in that a majority of linguistic writing problems that arise during text formulation are ill-structured (Heine, 2010; Roca de Larios et al., 1996). Some models of reference resource consultation sequences have appeared in previous studies (Kennedy & Miceli, 2001; Park, 2012). Although they were not developed from the problem solving perspective, they fit quite well the general problem solving cycle described above. Figure 4.1 juxtaposes these sequences with the problem solving cycle.

![Figure 4.1](image)

*Figure 4.1. Reference resource consultation models in previous studies.*

As can be seen above, the two corpus consultation models have roughly the same flow with each having steps that combine multiple steps in the problem-solving cycle. While Kennedy and Miceli’s model shows more clearly what the early steps in corpus consultation entail, Park’s reflects the reiterative nature of reference resource consultation by having the “Refine the query” step in it. Both models, when combined, thus serve as a good starting point for developing a model of i-Conc consultation by filling up each other’s missing steps. As can be visualized from Figure 4.1, when one model is superimposed onto the other, a full problem-solving cycle emerges.
Using these two models as a starting point, I developed a model of the typical i-Conc consultation process based on screen recording observations and the participants’ recollections of their i-Conc consultations. Despite my efforts to reduce steps to the broadest possible level, the resulting model (shown in Figure 4.2) is a little more complicated than the previous two models combined. This reflects the rather complicated interplay of the ill-structured nature of problems the participants dealt with and the use of a reference tool with multiple resources in the present research. As will be discussed in detail in later sections, many of problems the participants took up to solve using i-Conc pertain to (a) seeking English equivalents to Korean linguistic items or English forms that can express ideas conceived of in Korean, (b) collocates of specific words and phrases, and (c) alternative forms that can express the same intended meaning. These problems are often ill-structured and have multiple solutions or no clear solution with no explicit rules, principles, or criteria provided for reaching and evaluating the solution. This feature requires the problem solver to make judgments based on the given information at various points of problem solving. To add another layer of complexity, the multiplicity of reference resources featured in i-Conc presents multiple solution paths and this again requires the user to make judgments about the best resource for each query within a problem space.

![Figure 4.2. i-Conc consultation sequence.](image-url)
Each step in the process, though given a simple label, involves various operations and potentially a multitude of decisions. In the first step, the writer takes up a problem she encounters while writing and decides to consult *i-Conc* to find its solution. Here the writer’s judgment determines whether the given problem is soluble by consulting *i-Conc*. In the next step, the writer defines the problem by (a) determining whether the problem is about confirming existing knowledge or acquiring new knowledge (i.e., confirmatory or compensatory), and (b) zeroing in on the exact content of the problem, for example, whether it is a problem of finding a collocate, L2 equivalent, or alternative. At the same time the writer represents the problem by formulating a concrete question about the problem in either open or closed form (i.e., elicitation or verification). In the next step, the writer decides which resource to use and which option and query syntax, if necessary, to use in that resource, and what to actually type in the search box of *i-Conc*—i.e., which query term. Once a search strategy is formulated, the writer performs the query by selecting the resource and option, and then typing in the query term in the appropriate query syntax required in the given resource. When *i-Conc* returns query results, the writer evaluates whether these results provide an answer to the question she posed at the beginning of the consultation.

From this step downward, the process becomes more complicated as the problem solver needs to make a series of judgments and decisions and therefore multiple paths can emerge from each step. One of the most likely paths is that the writer finds a clear answer to the question from the initial query results. She then may stop there and apply the solution she has found to her writing. However, she may also perform another query to confirm whether it is the right one by, for example, consulting a different resource. If query results returned by a different resource corroborate, she then gets reassurance about the accuracy or appropriacy of the solution she has found and can apply it to her writing with greater confidence. Another likely path involves the answer the writer finds from the initial query results being not clear enough or only partial due to, for example, only a small number of instances. The writer then either applies this not-fully-satisfactory solution to her writing or performs another query that builds on and complements the initial one to get a more complete or satisfactory solution. In contrast to the two solution paths above, a third likely path is that the writer finds no answer from the initial query results due to difficulty in teasing out relevant answers from the results or no instances of the target
item searched for. The writer can then go back to the early steps, refine the definition or representation of the problem or search strategy and perform a new query. Or she just gives up performing any further queries and either attempts to solve the problem without the help of i-Conc or abandons the text formulation or revision of the given ideational content altogether. As seen in Figure 4.2, this part of the sequence forms a loop in which the number of queries the participant performs within a problem space depends on how long the participant is willing to persist at solving the given problem, and how much effort she puts into it. Therefore, the i-Conc consultation cycle can end with just a single query but can also iterate the loop multiple times.

As can be seen above, queries carried out in the cycle can be broken into two categories: *initial* and *subsequent* queries. As every problem space has an initial query, whether it is a single- or multiple-query consultation, and it is the initial query that reveals the participant’s motivation that initiates the problem solving cycle, I consider the initial query *primary* relative to subsequent queries that follow in a given problem space. The subsequent queries in turn are considered *secondary* to the initial query. In some discussions that follow, thus, the words *initial* and *subsequent* will be interchangeably used with *primary* and *secondary* respectively.

It should be noted that this i-Conc consultation sequence is a model that has been abstracted to cover many different i-Conc consultation situations, based on my empirical data, and that when one watches actual screen recordings, the participants may seem to skip and combine some steps in the model. Particularly in the early steps of the cycle, the participants often did not define or fully represent a problem nor formulate a strategy explicitly. They just went through these steps most of the time very quickly in the matter of a few seconds. This consultation sequence has been drawn from the participants’ data after the fact; the participants may not have been consciously aware of themselves going through all steps while engaging in problem solving using i-Conc.

Figure 4.3 provides an example of an i-Conc consultation that follows each step in the model. The example used here (YP11) is the same one used for explaining the concept of problem space in the previous chapter (and also see Appendix F).

<table>
<thead>
<tr>
<th>Step</th>
<th>Operations and decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encounter a problem</td>
<td>Initial state: … more <em>devastating images</em> and videos of youth culture and life have been produced and shared online …</td>
</tr>
<tr>
<td>Define and represent</td>
<td>Do native speakers use the phrase <em>devastating images</em> in the way I</td>
</tr>
</tbody>
</table>
the problem used it in the given context?

<table>
<thead>
<tr>
<th>Devise a search strategy</th>
<th>Resource:</th>
<th>Google</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query term:</td>
<td>devastating images</td>
<td></td>
</tr>
<tr>
<td>Query syntax:</td>
<td>Use ““ to get exact matches</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perform the query (Query 1, initial query)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click the Google tab of the i-Conc and type in “devastating images” in the search box</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluate the query results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query results:</td>
</tr>
<tr>
<td>- … devastating images of Texas’ drought …</td>
</tr>
<tr>
<td>- … devastating images of eco-disasters …</td>
</tr>
<tr>
<td>…</td>
</tr>
<tr>
<td>Evaluation:</td>
</tr>
<tr>
<td>I intended to mean something like images that make people uncomfortable, or irritated but the results showed that “devastating images” are mostly concerned with natural disasters. So I decided to perform another query.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 2 (subsequent query)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Devise a search strategy</td>
</tr>
<tr>
<td>Query term:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluate the query results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query results:</td>
</tr>
<tr>
<td>- … broadcast the devastating images and angry reaction from the scene…</td>
</tr>
<tr>
<td>Evaluation:</td>
</tr>
<tr>
<td>Here too the example showed that it is more about physical destruction or a disaster and there was only one instance. So I decided to look for a word that has a meaning closer to what I intended to say.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 3 (subsequent query)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define and represent the problem</td>
</tr>
<tr>
<td>What are the words that are similar in meaning to devastating?</td>
</tr>
</tbody>
</table>

| Devise a search strategy    | Resource: | COCA |
| Query term:                 | devastating |
| Query syntax:               | [=devastating] (*[=word] is a query operator used in COCA that returns the synonyms of the word) |

<table>
<thead>
<tr>
<th>Evaluate the query results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query results:</td>
</tr>
<tr>
<td>- overwhelming, disturbing, shocking, harmful, destructive, …</td>
</tr>
<tr>
<td>Evaluation:</td>
</tr>
<tr>
<td>The word disturbing seemed the closest to what I intended to say. So I wanted to search for disturbing image as the next query.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query 4 (subsequent query)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define and represent the problem</td>
</tr>
<tr>
<td>Do native speakers use disturbing images in the context I intended?</td>
</tr>
</tbody>
</table>

| Devise a search strategy    | Resource: | COCA |
| Query term:                 | disturbing image |

<table>
<thead>
<tr>
<th>Evaluate the query results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query results:</td>
</tr>
<tr>
<td>- Rated PG-13 for intense violent thematic material and disturbing images…</td>
</tr>
<tr>
<td>- … included several disturbing images of women being victimized by…</td>
</tr>
<tr>
<td>…</td>
</tr>
<tr>
<td>Evaluation:</td>
</tr>
<tr>
<td>This matched what I intended to mean. And there were also many more instances than devastating images. So I decided to replace devastating with disturbing.</td>
</tr>
</tbody>
</table>
Apply the solution found to the writing

Goal state: … more disturbing images and videos of youth culture and life have been produced and shared online …

Figure 4.3. An example of a full i-Conc consultation sequence

4.2.2 Types of problems and queries

This section presents descriptive statistics on the participants’ use of i-Conc during their writing assignment, mainly on problems types, and query types and purposes, and their breakdown across and within individual resources.

As noted in Chapter 3, the participants were encouraged to consult i-Conc as often as possible to familiarize themselves with the reference suite and had opportunities to develop their own query strategies. Some participants indeed used i-Conc very frequently even for their referencing needs that were not related to the writing assignment they chose for the present study. Therefore, not all the queries performed by the participants and recorded in the query log have been analyzed. Only those queries screen recorded and then recalled have been analyzed as the accurate identification of query functions and purposes was only possible by triangulating screen recording observations and the participants’ own recollections.

The participants as a group performed 3170 queries in total during their participation periods, which ranged from seven weeks to as long as 20 weeks depending on the individual (see Table 3.1 for each individual’s participation period). The actual number of queries done by the participants during these periods was much larger as only the queries entered in the search box of i-Conc were captured in the query log.

Table 4.3

Number of Queries Performed by the Participants

<table>
<thead>
<tr>
<th>Total queries</th>
<th>Jae</th>
<th>Yumee</th>
<th>Goeun</th>
<th>Jinho</th>
<th>Shia</th>
<th>Ian</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screen recorded queries (SR1-4)</td>
<td>227</td>
<td>1739</td>
<td>227</td>
<td>312</td>
<td>160</td>
<td>455</td>
<td>3170</td>
</tr>
<tr>
<td>Queries during the main task (SR2-4)</td>
<td>136</td>
<td>206</td>
<td>96</td>
<td>125</td>
<td>98</td>
<td>108</td>
<td>769</td>
</tr>
<tr>
<td>Total queries</td>
<td>98</td>
<td>121</td>
<td>52</td>
<td>93</td>
<td>72</td>
<td>79</td>
<td>515</td>
</tr>
</tbody>
</table>
Among these queries, 769 queries were screen recorded. Of those 769, 254 queries were recorded during the first screen recording (SR1) which was conducted as a practice session. The remaining 515 queries were recorded at three separate times (SR2, 3, 4) while the participants worked on the writing assignment they had chosen for the study. Table 4.3 above presents these numbers by participant.

Findings that follow are from the queries the participants screen-recorded while completing their self-chosen papers or the queries from SR2, 3, and 4.

4.2.2.1 Analysis of problems

The total number of screen-recorded queries while the participants completed their self-chosen authentic academic papers was 515. Through these queries they tried to solve 245 problems performing an average of 2.1 queries per problem space. Slightly less than 60% of the 245 problems, or 143, were of the confirmatory type (Conf)—i.e., cases in which the writer had a specific linguistic item in mind or in writing for a meaning she intended to express and tried to confirm whether they match accurately and/or appropriately in a given context—while 103 were of the compensatory type (Comp),—i.e., cases where the writer did not have a clear idea about how to convert into text a meaning she intended to express and so tried to extract potential target items from the reference resources. One problem was identified as both Conf and Comp and therefore the sum of problems of both types is one larger than the total number of problems. Of the two types of problems, Conf ones had an average of 2.22 queries within their problem spaces, slightly more than the 1.96 queries of Comp problems.

Then, what problems did the participants notice and attempt to solve using i-Conc? As noted in the earlier section, the motivation that initiated a problem space can be inferred from the purposes of the initial query within each problem space. The three most frequent purposes for which the participants performed their initial queries made up more than 50% of all of their purposes. The top three were finding or checking a collocate of a given word or phrase (Coll, 21.6%), finding an English equivalent (Equi, 17.1%), and confirming the presence or frequency of a given word or phrase (simple confirmation, or SC, 16.3%). A detailed description of queries for each of these main purposes is provided later in the section. These purposes were closely followed by checking if a given item conveys the intended meaning (Mn, 15.9%), finding a better alternative to a given item (uAlt, 11.0%), and checking the register of a given item (Reg,
6.1%). To break them down by problem type, the participants consulted *i-Conc* for the purposes of SC, Mn, and Coll most frequently for Conf problems while Equi, uAlt, and Coll for Comp problems. The graphs in Figure 4.4 show the proportions of the main initial query purposes within the total problems and within each problem type. (See Table 3.8 in Chapter 3 for detailed descriptions of what each code represents.)

![Figure 4.4. Frequencies of initial query purposes.](image)

### 4.2.2.2 Analysis of queries

I presented so far findings at the problem level and related to the purposes of initial queries. To extend the analysis of query purposes to include subsequent (or secondary) queries within the problem spaces, a somewhat different picture emerges. The three most frequent query purposes were SC (24.3%), Mn (21.0%), and Coll (18.1%), showing a different pattern from the initial queries, in which Coll and Equi were the two most frequent purposes. This reflects the participants’ tendency to confirm the presence/frequency of a given item (SC) retrieved from the initial query results or to check whether it matched what they intended to say (Mn). Figure 4.5...
shows the percentages of different query purposes in the main task data. A considerable portion (13.2%) of the total queries had more than one purpose and therefore the sum of percentages of individual query purpose categories exceeds 100%.

![Figure 4.5. Frequencies of query purposes.](image)

To look in greater detail at the major query purposes, first, SC queries, the most frequent queries in the data, were by and large performed when the participants were not confident about a phrase or a structure that they came up with and tried to check whether it was acceptable to use the item in the given context. The participants appeared to have had some awareness that a text cannot be formulated by simply connecting a series of single words according to syntactic rules but that much part of connected text is formulaic, and behaves as chunks, which was also explicitly explained during the training sessions. When they formulated a multi-word string that seemed novel or not conventional, they asked themselves whether it sounded natural, whether native speakers would normally use it in the way they used it, and they tried to get answers through SC queries. The target items searched for through SC queries were mostly word sequences ranging from noun phrases (e.g., mixed tone, YP50-1) to long word strings that are not syntactically complete (e.g., was taken as an insulting accusation, SP32-3). Being more than a single word in general, SC queries were longer than other types of queries. An SC query had an average of 3.29 words while the average number of words per query for all query purposes was 2.06.

Queries for checking meanings (Mn), the second most frequent query purpose, were mostly carried out when the participants were not sure about whether an item actually meant
what they thought it meant. That is, the participants already had a rough idea what the item meant and attempted to verify whether it was appropriate in the given context. Therefore, Mn queries performed by the participants went much further than simply checking the definition of a word or phrase, to include, in many cases, perusing lexico-grammatical environments surrounding the items. Although the participants still used bilingual and monolingual dictionaries for single words, which made up a vast majority of Mn queries (68.5%), they also often verified the match between multi-word target items and their intended meanings through checking the contexts provided in the concordance lines in corpus-based resources such as COCA and JTW.

Finding (or checking) collocates (Coll) was the third most frequent purpose for which the participants consulted i-Conc. As presented earlier, however, Coll was the most frequent purpose of the initial (primary) queries, and Coll queries were performed more as primary (55.8%) than as secondary queries (44.2%). By contrast, SC and Mn, the two most frequent query purposes, were used much more frequently as secondary than as primary queries within a problem space. A typical Coll query was performed when the participant had an item in mind (a node word/phrase) and attempted to find or confirm a collocate in a certain word class that typically or frequently co-occurred with the node to deliver the intended meaning. The most frequent collocate type searched for in terms of word class was preposition (34 out of the total 95 queries, or 35.8%), followed by verb (27.4%), adverb (12.6%), and adjective (12.6%). When looked at in terms of the relationship between the node and the collocate, verb collocates of noun objects (V + n) were the most frequently searched for (18.9%). Following this were preposition collocates that occur after nouns (n + P) (14.7%) and adjective collocates that modify nouns (Aj + n) (9.5%).

Lastly, finding L2 equivalents (Equi) was the fourth most frequent purpose for which the participants consulted i-Conc (10.9% of all queries). What set Equi queries apart from other queries is obviously that they were all in Korean (with only one exception) and were performed, as expected, when the participants could not (or did not) make any hypotheses about L2 forms for an intended meaning, which was the case with most queries in English, and instead they sought suggestions from the resources. Of particular note is that Equi queries were especially performed when the participants wrote sections discussing their personal experiences, emotions, and things of a more narrative nature, most of which they experienced in their L1. Some other
common characteristics of Equi queries were that they were mostly single-word queries (1.25 words on average), and performed as the initial queries of compensatory problems (42 out of the total 56, 75%). Table 4.4 summarizes what has been discussed about the major query purposes above.

Table 4.4

<table>
<thead>
<tr>
<th>Major Query Purposes</th>
<th>N=245</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>f</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Simple confirmation (SC)</td>
<td>125</td>
</tr>
<tr>
<td>Intended meaning (Mn)</td>
<td>108</td>
</tr>
<tr>
<td>Collocation (Coll)</td>
<td>95</td>
</tr>
<tr>
<td>L2 Equivalent (Equi)</td>
<td>56</td>
</tr>
</tbody>
</table>

*Note.* a. The ratio of primary queries to secondary queries.

As described in Chapter 3, queries were also classified into two types depending on the form of the inquiry: elicitation and verification. An elicitation query was performed to find answers to a *what* or *how* question (an open form question), and usually returned more than one choice in the query results. A verification query was performed to find answers to a *whether* question (a closed form question) and to “verify” from the query results the presence/frequency, accuracy or appropriacy of the target item in a given context. Thus, a compensatory problem solving cycle usually started with an elicitation query to extract from i-Conc what was unknown to the problem-solver while a confirmatory problem solving cycle with a verification query was usually to test a hypothesis made based on the problem-solver’s existing knowledge. However, this was not an absolute rule as some participants strategically started with an elicitation query to solve a confirmatory problem. This tendency is discussed in greater detail in later sections. Figure 4.6 illustrates a typical query of each type.
About 60% of the participants’ queries were of a verification type while 40% were of an elicitation type. This allocation suggests that the participants consulted i-Conc more to “check” their hypotheses about L2 forms by asking a yes/no question than to “explore” by posing an open question that extracted potential candidates for the target items. More specifically, verification queries were performed for the purposes of SC, Mn, Reg (register), and Coll, to list them in order of frequency while elicitation queries for Coll, Equi, uAlt (upgrading alternative) and sAlt (simple alternative) (see Table 3.8 in Chapter 3 for descriptions for query purposes). As can be seen above, each type was performed for a unique set of purposes respectively while sharing some purposes such as Coll. As for Coll queries, 70.8% were elicitation queries while 29.2% were verification queries.

Finally, queries were classified into four linguistic categories: lexical, lexico-grammatical, grammatical, and stylistic categories (see Table 3.8 in Chapter 3 for definitions of the categories). The vast majority of participants’ i-Conc consultations were concerned with lexical and lexico-grammatical issues (45.6% and 43.7% respectively). Queries for stylistic and
grammatical issues took up 11.7% and 8.3% respectively. That grammatical issues ranked lowest may reflect the high language proficiency levels of the participants, who overall had a great control over syntactic rules, judging from screen recording observations.

The queries classified into each linguistic category, to a large extent, had a different set of purposes (see Table 4.5). Of particular note is that the participants addressed 66% of the lexical issues using the more traditional reference resources—i.e., the resources in the Dictionary tab of i-Conc—while using the concordancers for 85.8% of the lexico-grammatical and 91.7% of the stylistic issues. This strong tendency to use different types of resources for different purposes and linguistic issues demonstrates that the participants were in general well aware of when and for what to use each type of resource. Table 4.5 summarizes the types and linguistic categories of the participants’ queries presented above.

Table 4.5

<table>
<thead>
<tr>
<th>Types and Linguistic Categories of the Participants’ Queries</th>
<th>N=515</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query</td>
<td>f</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Functional type</td>
<td></td>
</tr>
<tr>
<td>Verification</td>
<td>317</td>
</tr>
<tr>
<td>Elicitation</td>
<td>209</td>
</tr>
<tr>
<td>Linguistic category</td>
<td></td>
</tr>
<tr>
<td>Lexical</td>
<td>235</td>
</tr>
<tr>
<td>Lexico-grammatical</td>
<td>225</td>
</tr>
<tr>
<td>Stylistic</td>
<td>60</td>
</tr>
<tr>
<td>Grammatical</td>
<td>43</td>
</tr>
</tbody>
</table>

Note. The sum of queries of different types, and also the sum of queries in different linguistic categories are both larger than the total number of queries, 515. That’s because there are queries classified into more than one type or linguistic category.

7 Again, the percentage numbers for these linguistic categories, when combined, exceed 100% since a considerable amount of the participants’ queries were classified into more than one category.
4.2.3 Resources consulted

In this section I present findings on the consultation frequency of each resource (type) in i-Conc, and their distributions across query purposes, types and linguistic categories. To start with consultation frequencies, the bilingual dictionary (Bilingual) was the most frequently consulted resource in i-Conc (25.6% of all queries), which was closely followed by COCA (25.4%). Google was a relatively close third (18.1%), followed by JTW (11.1%), Monolingual (7.8%), and Thesaurus (5.2%). CSE and GS ranked lowest, consulted for less than 5% of the total queries respectively. To group the resources into concordancers and dictionaries, the participants performed 61.4% of queries using the concordancer-type resources compared to 38.6% using the dictionary-type resources.

Turning to the purposes i-Conc was consulted for, Figure 4.7 provides consultation frequencies of the eight different resources featured in i-Conc and the purposes for which each resource was consulted. The most striking patterns emerging from Figure 4.7 are: (a) the resources in general had a unique set of dominant purposes for which they were consulted, and (b) though it was not an absolute trend, the more frequently consulted a resource was, the wider the range of purposes for which that resource was consulted.

The top purpose for which Bilingual was consulted was finding L2 equivalents (Equi) (41.7%). As can be seen in Figure 4.7, Equi queries were performed almost exclusively on Bilingual. Combined with the fact that Equi was one of the most frequent purposes for which the participants consulted i-Conc, Bilingual being an exclusive tool for Equi queries contributed to its overall frequent use. However, that was not the only contributor. As shown in the length of the Bilingual column in Figure 4.7, it was consulted for the widest range of purposes of all resources. This should be attributed in large part to the versatility of the specific online bilingual dictionary featured in i-Conc—i.e., Naver online Korean-English dictionary (Naver in short). Unlike paper dictionaries, Naver has evolved rapidly over the last few years adding many source dictionaries, English-Korean translation pairs of the sentence level and above, and new
Note: As in the previous tables and charts, the sum of the frequencies of each purpose exceeds the total number of queries done in a resource as a considerable portion of the queries had more than one purpose. In those cases, the queries were counted more than once.

**Figure 4.7.** Frequencies of query purposes, types and linguistic categories in each resource.
reference features taking advantage of its huge online capacity for information processing and storage. Now it has become more of dictionary, thesaurus, and search engine rolled into one. One of its features the participants found most useful was what can be seen as a primitive form of a Korean-English parallel corpus. For a Korean query—be it a word, phrase or even sentence—it returns results (Korean and English sentences in parallel) in the concordance format along with the usual dictionary information of each constituent word. The participants took full advantage of these features and used the resource for such purposes as SC (7.6%), sAlt (6.1%), and uAlt (4.5%) that are typically met with a large corpus and thesaurus.

On the concordancer side, COCA proved to be the most frequently used concordancer resource, consulted almost as often as Bilingual. Reflecting its multiplicity of features and sophisticated query options, COCA was used for a wide variety of purposes, which were distributed relatively evenly in terms of percentage.

Google, the third most frequently used resource, was the most salient in its pattern of having a predominant purpose for which it was consulted. Almost 80% of the time, it was used for SC purposes. This tendency can be attributed to its nature as a commercial search engine, as was discussed in previous chapters. Google searches the huge amount of language data available on the Web, far larger than other corpus tools, and therefore lends itself well to SC queries, as it returns a considerable number of instances of the searched-for items, even multi-word sequences, and relatively rare or idiosyncratic phraseologies. However, due to the very nature of Web-based language data, which is a mélange of different genres, registers, and provenances, the participants did not consult Google for other purposes as often.

JTW was consulted for Coll nearly 60% of the time. Given that JTW is a concordancer specifically designed for word combination searches, this result should not come as a surprise. Citing the way it displays query results and its simplicity of use (e.g., requiring no special query syntax), some participants rated this resource as the best concordancer in i-Conc. However, during the study period, the server crashed quite often, and even when it wasn’t down, the loading of results took so long that one participant (Yumee) stopped using it after a few trials. In addition, the source corpus this resource draws on is relatively small and so its query results sometimes missed relatively frequent word combinations. This caused some participants (Jae, and Shia) to give up searching the resource quite often. Otherwise, the resources would have been consulted more and the overall results might have been different.
Monolingual and Thesaurus were consulted for the purposes usually expected for them. Although they are all provided online, they differ from Bilingual in that they offer no features that go beyond those provided in their paper versions. Therefore, Monolingual was consulted for no purposes other than checking intended meanings (Mn, 70%) while Thesaurus was used for finding alternatives (sAlt, and uAlt, 81.4%).

Finally, CSE and GS ranked lowest, each being consulted less than 5% of the time. As with the greater-than-expected use of Bilingual discussed above, this underuse of the resources by the participants comes as a surprise. These two resources were included in the reference suite specifically to meet the needs expected to arise during disciplinary writing by graduate students. However, after several attempts to use these two in the beginning of the study, the participants found them limited due to their small corpus size and many incidences of returning no matches of the queried items. When they consulted these two resources, however, the participants did so for the purposes of Reg (register check) and SC, which were what these resources specifically designed for.

What has been discussed so far suggests that the participants were largely aware of what each resource is effective at, and they used the resources in i-Conc for the purposes for which they were intended. However, the resource data grouped into query purposes will provide another angle from which to see how the participants chose different (types of) resources for different purposes.

<table>
<thead>
<tr>
<th>Query purpose (the total number)</th>
<th>SC (125)</th>
<th>Mn (108)</th>
<th>Coll (95)</th>
<th>Equi (56)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>%</td>
<td>f</td>
<td>%</td>
</tr>
<tr>
<td>Google</td>
<td>72</td>
<td>57.6</td>
<td>36</td>
<td>33.3</td>
</tr>
<tr>
<td>COCA</td>
<td>19</td>
<td>15.2</td>
<td>28</td>
<td>25.9</td>
</tr>
<tr>
<td>Biling.</td>
<td>10</td>
<td>8.0</td>
<td>27</td>
<td>25.0</td>
</tr>
<tr>
<td>CSE</td>
<td>10</td>
<td>8.0</td>
<td>9</td>
<td>8.3</td>
</tr>
<tr>
<td>GS</td>
<td>10</td>
<td>8.0</td>
<td>Th</td>
<td>5</td>
</tr>
<tr>
<td>JTW</td>
<td>2</td>
<td>1.6</td>
<td>CSE</td>
<td>2</td>
</tr>
<tr>
<td>Monoling.</td>
<td>2</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>uAlt (43)</th>
<th>Reg (39)</th>
<th>sAlt (30)</th>
<th>GU (23)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>%</td>
<td>f</td>
</tr>
<tr>
<td>COCA</td>
<td>19</td>
<td>44.2</td>
<td>15</td>
</tr>
<tr>
<td>JTW</td>
<td>9</td>
<td>20.9</td>
<td>10</td>
</tr>
<tr>
<td>Thesaurus</td>
<td>9</td>
<td>20.9</td>
<td>9</td>
</tr>
<tr>
<td>Biling.</td>
<td>6</td>
<td>14.0</td>
<td>Google 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Biling. 2</td>
</tr>
</tbody>
</table>

Figure 4.8. Frequencies of resource use for major query purposes.
Figure 4.8 provides the frequencies of resource use for major query purposes. Firstly, it clearly shows that within each query purpose, there is largely one predominant resource or resource type that makes up more than half of all queries. While for SC, Coll, uAlt, and Reg the concordancer-type resources were used predominantly, the dictionary-type resources were the dominant resources for Mn, Equi, sAlt and GU (general usage). However, within the same types, different subtypes showed differing dominance in those purposes. For example, while Google was the main resource for SC, COCA and JTW were the dominant resources for Coll, and uAlt. This trend can be attributed to the different nature of the concordancers, that is, differences between the untagged and “raw” corpus tools (Google, GS, and CSE) and the tagged and mediated corpus tools (COCA, and JTW) that are amenable to more sophisticated searches (Frankenberg-Garcia, 2005). Likewise, within the dictionary-type resources, each resource proved to be the top resource for the purposes they are normally expected to serve (Bilingual for Equi, and Bilingual and Monolingual for Mn). Secondly, as seen in Figure 4.8, COCA comprised, more or less, the major proportion of all major query purposes except in Equi. Combined with the findings presented earlier—that COCA was used equally frequently for elicitation and verification and that it is the most sophisticated concordancer—this proves COCA is a balanced resource that can be consulted for various purposes and functions.

Turning now to the distributions of query types and linguistic categories within each resource, as can be seen in Figure 4.7, a few salient patterns stand out. First, the two query types are relatively evenly distributed in Bilingual, COCA, and JTW while the verification type is by far dominant in relation to the rest of the resources, reflecting the different major purposes each resource serves. This pattern again shows clearly as with the query purposes that within the same resource type (i.e., concordancers and dictionaries) resources do not necessarily share functionalities and uses to which they can be put. For instance, reflecting the differing purposes they were consulted for, the untagged and “raw” corpora (Google, GS and CSE) were predominantly consulted for verification queries (over 90%) while the tagged and cleaned corpora (COCA and JTW) were used evenly for both query types. Second, however, as presented earlier, the two resource types were distinct from each other in their query distributions in terms of linguistic categories. The dictionary-type resources were mostly
consulted for queries about lexical concerns while the concordancer type was consulted mostly for queries about lexico-grammatical and stylistic concerns.

4.3 Effects of using i-Conc on problem solving performance

In this section, I present findings regarding the effects of using i-Conc on problem solving performance. The effects on problem-solving performance here specifically refer to the extent to which i-Conc consultations led to the successful resolution of the problems the participants noticed and chose to address using the reference suite. As described in the previous chapter, each problem space was coded on the following three dimensions: (a) whether the participant was satisfied with the consultation results—i.e., whether they perceived as satisfactory the answer they found to the given problem, (b) whether the participant actually found a correct solution to the given problem, and correctly applied it to their writing, and (c) whether the participant abandoned the consultation (see Section 3.4.3 in Chapter 3 for detailed descriptions of the coding scheme). Each problem space was then given a set of codes that consists of “S” (satisfied) or “D” (dissatisfied) for (a); “C” (correct), “I” (incorrect), or “N/A” for (b); and “A” (abandoned) or blank for (c). Depending on the set of codes a problem space was assigned, it was classified into one of seven groups labelled A to G. Appendix G shows the set of codes for each group and the solution path and problem solving result each set of codes represents.

Table 4.6 presents the numbers and percentages of the problem spaces belonging to each group. Different combinations of these groups provide multi-level findings. In the following two subsections, I present a brief overview of findings different combinations of groups provided and then more specific findings from each combination of groups.

---

8 The final drafts of the participants’ writing assignments still contained varying amounts of lexical and grammatical mistakes and stylistic infelicities that had been either unnoticed or noticed but not addressed with i-Conc. Thus it should be emphasized that specific figures of successful problem solving to be presented in the following sections should not be taken to represent the overall quality of the participants’ papers.
Table 4.6

*The Number and Percentage of Problem Spaces Belonging to Each Group*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>158</td>
<td>13</td>
<td>20</td>
<td>10</td>
<td>32</td>
<td>4</td>
<td>8</td>
<td>245</td>
</tr>
<tr>
<td>%</td>
<td>64.1</td>
<td>5.3</td>
<td>8.6</td>
<td>4.1</td>
<td>13.1</td>
<td>1.6</td>
<td>3.3</td>
<td>100</td>
</tr>
</tbody>
</table>

4.3.1 Brief overview of different code sets

Groups A and B in Table 4.6 represent the problem spaces where *i*-Conc consultation resulted in a correct text formulation or revision of the target item, and thus had a positive effect on the participant’s writing. Groups C and D represent the problem spaces where *i*-Conc consultation had a negative effect and resulted in an incorrect text formulation or revision.

The participants’ different perceptions of their own *i*-Conc consultation results can also be drawn from Table 4.6. Groups A and C combined show the proportion of the consultations the participants were satisfied with, that is, which they perceived as successful. The remaining groups (B, D, E, F, and G) combined represent *i*-Conc consultations the participants were dissatisfied with for different reasons. The difference between B and D on the one hand, and E, F, and G on the other, is that in the former the participants, although dissatisfied with the consultation results, retrieved something from the results and applied it to writing while in the latter, they abandoned the consultations.

Finally, as can be seen above, abandoning *i*-Conc consultation did not necessarily mean that the participants gave up the problem solving. Groups E and F represent the cases where the participants reached a solution on their own after their initial effort to use the reference suite did not lead to a satisfactory solution. Thus Groups A, C, and E here represent the problem spaces that led to a successful resolution of the problem with or without the help of *i*-Conc while the rest represents unsuccessful problem solving cases.

Merging the groups as discussed above and giving them more intuitive labels produces Table 4.7.
Table 4.7

Classification of Problem Spaces According to their Problem Solving Results

<table>
<thead>
<tr>
<th></th>
<th>Positive effect</th>
<th>Negative effect</th>
<th>Satisfied consultation</th>
<th>Dissatisfied consultation</th>
<th>Successful problem solving</th>
<th>Unsuccessful problem solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f )</td>
<td>170</td>
<td>31</td>
<td>178</td>
<td>67</td>
<td>202</td>
<td>43</td>
</tr>
<tr>
<td>%</td>
<td>69.4</td>
<td>12.7</td>
<td>72.7</td>
<td>27.3</td>
<td>82.5</td>
<td>17.5</td>
</tr>
</tbody>
</table>

4.3.2 Findings from each combination of groups

I begin with positive and negative effects of using \( i \)-Conc on the participants’ writing. As can be seen Table 4.7, out of 201 problem spaces where a solution was retrieved through \( i \)-Conc consultation and applied to the participants’ writing, 170 or 69.4% resulted in a correct text formulation or revision. That is, nearly 70% of the time, \( i \)-Conc consultation had a positive effect on the participant’s problem solving. On the other hand, \( i \)-Conc use led to an incorrect text, having a negative effect on writing, in 12.7% of the problem spaces.

Tables 4.6, and 4.7 show that the actual results, that is, positive and negative effects, did not necessarily match the participants’ perceptions of the consultation results. The participants were dissatisfied with 13 out of the total 170 problem spaces where \( i \)-Conc consultation had a positive effect whereas 20 problem spaces ended up with an incorrect text formulation or revision even though the participants perceived them as successfully solved. These cases were caused by different factors such as a wrong solution retrieved from the query results or misapplication of the chosen solution to writing.

In terms of problem type, the rate of positive effects by \( i \)-Conc was considerably higher with Conf problems. Out of a total of 143 Conf problems, 109 or 76.2% were successfully solved through \( i \)-Conc consultation while in 61 out of 103 Comp problems (59.2%), a correct solution was found and successfully applied to writing. This pattern suggests that overall the participants had greater difficulty in solving Comp problems using \( i \)-Conc.

To look at the primary query purpose of (or initial motivation for) each problem, AP (argument pattern) emerged on top (90.9%) in the rating of positive effects by \( i \)-Conc among the major purposes that were consulted more than five times by the participants. It was followed by NE (noun endings), Mn, Coll, and GU. At the opposite end, sAlt (simple alternatives), uAlt, and Reg proved to be the areas where \( i \)-Conc consultation had the least positive effects (see Table
4.8). There are two patterns worth noting from Table 4.8. First, the purposes that could normally be met through a simple query or a short series of queries ranked higher than those that were usually met through multiple query refinements (such as Equi, and SC). Second, the purposes coming in at the bottom of the positive effect rate ranking (such as sAlt, uAlt, and Reg) showed the highest rates of abandoned consultations. The second pattern suggests that these purposes may be the hardest to find satisfactory solutions for but that these may also be ones that were readily given up on when faced with dissatisfying query results. Indeed, checking registers (Reg) or finding alternatives (sAlt) are all stylistic problems in which the writer already has L2 forms to fall back on even if she fails to get answers from i-Conc. These problems may have been judged not as important or urgent as other problems for which they did not have L2 forms that participants could use with confidence for their intended meanings.

Table 4.8
Effects of i-Conc Consultation by Problem Type and Primary Query Purpose

<table>
<thead>
<tr>
<th>Problem type</th>
<th>Total</th>
<th>Positive (Groups A, B)</th>
<th>Negative (C, D)</th>
<th>Abandoned (E, F, G)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f</td>
<td>%</td>
<td>f</td>
<td>%</td>
</tr>
<tr>
<td>Conf</td>
<td>143</td>
<td>109</td>
<td>19</td>
<td>13.3</td>
</tr>
<tr>
<td>Comp</td>
<td>103</td>
<td>61</td>
<td>13</td>
<td>12.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary query purpose</th>
<th>Total</th>
<th>Positive (Groups A, B)</th>
<th>Negative (C, D)</th>
<th>Abandoned (E, F, G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>11</td>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NE</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mn</td>
<td>38</td>
<td>30</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Coll</td>
<td>52</td>
<td>41</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>GU</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Equi</td>
<td>44</td>
<td>32</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>SC</td>
<td>40</td>
<td>29</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Art</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Reg</td>
<td>15</td>
<td>9</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>uAlt</td>
<td>27</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sAlt</td>
<td>15</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Note. a. Conf: confirmatory, Comp: compensatory

In a considerable number of problem spaces, the participants abandoned their i-Conc consultation midway due to difficulties finding relevant items from the initial query results (see
Groups E, F and G in Table 4.6). Out of 245 screen-recorded problems, the participants gave up i-Conc consultation in 44 problems, or 18.0%, midway and either found a solution on their own (Groups E and F), or abandoned the text formulation/revision of the given item altogether (Group G). Of note is that the participants still correctly solved on their own many of the problems for which they had given up i-Conc consultation. In 32 out of 44 problems, or 72.7%, a correct text resulted (Group E) while four, or 9.1%, ended up with an incorrect text (Group F). By problem type, the rate of abandoned consultations was higher with Comp problems (28.2%) than with Conf problems (10.5%). This trend adds further evidence for the finding presented earlier that the participants had greater difficulties solving Comp problems than Conf problems by consulting i-Conc (see Table 4.10).

4.4 Strategies and pitfalls in i-Conc consultation

As presented above, overall the participants used the reference suite effectively in solving problems they encountered during the writing assignment. While doing so, they used strategies at various levels that contributed to successful problem solving. Some of these strategies were taught during the tutoring sessions but a majority of them were developed by the participants themselves through repeated use of i-Conc. On the other hand, the participants‘ i-Conc use still proved to be ineffective and led to incorrect text formulation or revision in a considerable number of problems. This was due in part to a lack of language proficiency or the participants‘ ability to interpret the results and tease out relevant information needed for solving the problem at hand. However, a majority of the participants‘ unsuccessful problem solving cases can be traced back to a lack of reference consultation skills and discipline—i.e., pitfalls that could have been avoided.

In this section, I present some of the strategies and pitfalls identified from the triangulated data that contributed to the positive and negative effects and the participants‘ different perceptions of their i-Conc consultations. As with other findings presented in this chapter, these main strategies and pitfalls presented here were commonly observed in general across the participants.
4.4.1 Strategy use in i-Conc consultation

In the beginning of the study, as part of the training, the participants were provided with basic strategies for choosing the appropriate resources for different query purposes and for formulating query terms on each resource. In the earlier sections, the results showed that the participants effectively employed some of these strategies mostly by consulting the resources appropriate for their query purposes.

The strategies used by the participants are divided into query strategies and evaluation/application strategies. Query strategies here refer to strategies used largely in the first half of the i-Conc consultation cycle presented in Section 4.2.1. They were used for formulating or refining problem definition/representation and/or query term, especially when the initial query did not provide a clear or immediate solution to the given problem. Evaluation/application strategies were, in turn, used when determining what to choose from multiple alternatives presented in the results or whether an item was an appropriate solution to the given problem.

4.4.1.1 Query strategies

One of the most commonly used strategies was to use a different resource to expand or build on the previous query. More specifically, the participant got a partial or potential solution from an initial query she performed in a resource and then carried out another query using a different resource to get a more complete solution or corroboration. Or she performed a preliminary query in a resource to get a hint for more accurate and appropriate words or strings to be queried. As one of the highly encouraged reference resource consultation strategies in previous studies (e.g., Flowerdew, 2009; Frankenberg-Garcia, 2005; Kennedy & Miceli, 2010), this strategy takes full advantage of the multiplicity of resources featured in i-Conc and different strengths of each resource. The most common sequence in using this strategy was getting a hint from Bilingual for an L2 form to express the intended meaning and then performing an English query on COCA or Google for its acceptability or frequency.

When initial queries returned unsatisfactory results, one common strategy used by the participants was refining them by replacing a constituent word or structure with an alternative. In many problem spaces where the initial query on a word sequence returned no or few instances occurring in the context intended, the participants replaced one of the constituent
words with its synonyms. In other problem spaces, some participants changed the syntactic structure of the queried word sequence when it initially returned insufficient results. For example, when looking for an adjectival collocate of a given noun, one participant used a “noun + BE + adjective” structure as a query term after initially querying an “adjective + noun” structure. Through this lexical or syntactic replacement the participants often closed in on the target item they had set out to find.

Another common strategy used was making queries more general or specific when the original ones returned too few or too many instances of the searched-for item. This strategy was mainly used for SC queries where the number and specificity of words greatly affect the quantity of results. When a query returned no or few instances because it was too long and/or specific, the participants often reduced the number of words or replaced some words with more general ones or wild cards. When a query returned too many instances to go through because it was a single word or short phrase that was too broad and general in meaning, the participants then made the query more specific by adding more words or using words that indicated more specific referents.

If what has been described above can be seen as strategies for query refinement, some participants, mostly Jae and Ian, used a highly advanced and efficient strategy in their problem representations. As discussed in Chapter 3 and also in this chapter, Conf problem spaces usually start with a verification query, which represents the problem as a closed-form question (i.e., a yes/no question) about whether an linguistic item one has in mind or writing is correct or appropriate in the given context. If the query returns the results that say “no” to the question, the problem solver needs to formulate another query to continue to solve the given problem by refining the query and/or using a different resource. In some Conf problems spaces, however, some participants purposely started off with an elicitation query, normally used for Comp problem solving where the problem solver usually seeks to extract what is unknown to her. This strategy involves representing the problem in an open-form question—e.g., what is the most appropriate adverb that modifies the given verb?—even though the problem solver has a specific item (e.g., a specific adverb) in mind. Then an elicitation query typically returns multiple alternatives to choose from. If the initial item that the problem solver has in mind is included in these choices, then she gets reassurance that she can apply it to her writing or choose one that sounds better from the given alternatives. Even when the item is missing in the results,
she can still choose one she judges to be the best fit to the given context without having to perform further queries. For query purposes such as Coll and Equi, this strategy proved to be an efficient way of saving the problem solver performing extra queries.

Some participants utilized advanced search options that they discovered on their own in the individual resources as well. For instance, while using Google, Goeun sometimes confined her searches to a specific file type such as PDF to reduce “noise” caused by the nature of a commercial search engine. Jae ventured into search options that were not covered in the tutoring sessions and attempted creative ways to query target items, such as a query with two synonym brackets on COCA. Table 4.9 provides a brief description of each strategy discussed above and examples from the participants’ data.

Table 4.9

Main Query Strategies and Examples

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expanding a query using different resources</td>
<td>YP24 Initial query on COCA: <em>discussion [vb</em>] [j*] (it would return all the instances of “discussion + be + adjective”) Subsequent query on Google: “discussion was deep” Description: Yumee looked for a predicative adjective that modifies the noun “discussion” and ran the initial query as above. From the results she chose “deep” and ran another query on Google this time to confirm whether “discussion was deep” is acceptable.</td>
</tr>
<tr>
<td>Replacing a constituent word, POS, or syntactic structure</td>
<td>GP4 Initial query on Google: “from a Marxist perspective that” Subsequent query on Google: “from a Marxist perspective where” Goeun intended to check whether the string was used in her intended context where “perspective” is used as a shell noun but the initial query returned few instances of that. So she attempted another query with “that” replaced by “where”.</td>
</tr>
<tr>
<td>Making a query more general</td>
<td>JP20 Initial query on COCA: <em>statistics regarding</em> Subsequent query on COCA: <em>statistics [i</em>] (it would return all the instances of “statistics + preposition) Description: Jae intended to check whether “regarding” is an appropriate preposition coming after “statistics” but the initial query returned a relatively small number of instances. Then he refined the query more general by replacing “regarding” with [i*] so that the query would return all the prepositions that come after the word “statistics”.</td>
</tr>
<tr>
<td>Making a query more specific</td>
<td>YP34 Initial query on Google: <em>concerning themselves with</em> Subsequent query on Google: <em>concerning themselves with questions</em> Description: Yumee intended to check whether the phrase “concerning themselves with” can be acceptable in the given context but the query returned too many hits, so she made the query more specific by adding “questions”.</td>
</tr>
</tbody>
</table>
Starting a Conf PS with an elicitation query

**IP22-1**

*Query on JTW: limitation*

Description: Ian wanted to check whether the verb “overcome” is an appropriate verb collocate of “limitation”. Instead of querying “overcome [the] limitation”, he carried out an elicitation query on JTW that would display verbs that typically occur with “limitation”.

**JP66**

*Query on COCA: [=existing] legislation*

Description: Jae intended to check whether “existing” is a good collocate of “legislation”. But he also wanted to see other alternatives he can use if “existing” is not the best one.

Using advanced or creative search options

**JP35-4**

*Query on COCA: [=lose][=wage]*

Description: While struggling to find the appropriate L2 equivalent to 음금상실 (loss of income), he came up with the verb “lose” and the noun “wage”. Yet he was confident with neither of them and performed a query of two synonym brackets to see all the combinations of a verb in the same semantic group as “lose” and a noun in the same semantic group as “wage”.

4.4.1.2 Evaluation/application strategies

Though not as varied as query strategies, the participants used a number of strategies in evaluating query results and applying what they judged to be the solution to their writing. The most frequently used strategy in evaluating query results was checking the frequency of the searched-for item in the results—i.e., the number of tokens on COCA and JTW and the number of hits on Google. The participants tended to choose the most frequent one from different alternatives in elicitation queries while using frequency as a criterion for the acceptability of a queried item in the given context in verification queries. However, some participants took into account their own familiarity with alternatives and the authority of the source texts from which the query results were retrieved.

One major strategy the participants used in integrating the solution retrieved from the query results into their writing was modifying it for their intended contexts. In many, if not all, problem spaces, the participants did not accept uncritically what they saw as a solution from the results and apply it to their writing verbatim. Rather, they adjusted grammatical and semantic aspects to fit their own intended context.

Finally, another common strategy used was borrowing words and phrases found by chance discoveries while carrying out unrelated queries. Referred to by different names such as “serendipitous learning” and “discovery learning,” this behaviour has been indeed discussed as one of the major benefits of using corpus tools in the learner concordancing literature (e.g.,
Bernardini, 2004). Through these chance discoveries, the participants noticed problems that they had been unaware of.

Table 4.10 documents the evaluation/application strategies and their examples.

**Table 4.10**  
*Evaluation/application Strategies and Examples*

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Example</th>
</tr>
</thead>
</table>
| Choosing the most frequent one | JP60  
Query on COCA: *f=strongly influenced*  
Query results: *strongly influenced* (tokens: 318), *greatly influenced* (214), *deeply influenced* (97), *clearly influenced* (35) ...  
Application: *strongly influenced*  
Description: Jae performed a query to check whether “strongly” is an appropriate adverb that modifies “influenced” and if not, what alternatives there are to choose from. The query results showed “strongly” is the most frequent adverb in the given semantic group and he chose it. |
| Using familiarity with the item or authority of source as criteria | JP1  
Query on Google: *allintext: “globalizing South Korea”*  
Application: *in a globalizing South Korea*  
Description: Ian intended to check whether the phrase should come with the indefinite article or not and the results were mixed but he found instances of the phrase used with the indefinite article, which were mostly from an often-cited scholar. So he decided to use the indefinite article. |
| Contextualizing the retrieved item | JP2  
Query on COCA: *fr* regulated  
Query results: *highly regulated*, *heavily regulated*, *tightly regulated*, *federally regulated* ...  
Application: *nationally regulated*  
Description: Jae carried out a Coll query for adverb collocates that modify “regulated” and found “federally” the closest to the meaning he intended but realized that he could not directly use it as what he was discussing in his paper was about a government regulation in Korea. He then modified the *i-Conc* finding to fit the Korean context and used “nationally”. |
| Borrowing items discovered by chance | SP17  
Query on Google: *allintext: “systemic concealment”*  
Application: *reveals systemic concealment*  
Shia intended to confirm whether the phrase “systemic concealment” is used in the given context. While checking its instances, she found an instance where the verb “reveal” was used with the target item, and she replaced the original verb “exhibit” with “reveal”. |

**4.4.2 Pitfalls in *i-Conc* consultation**

Although the participants successfully solved many of the problems through *i-Conc* consultations using such strategies as discussed above, there were still a considerable number of cases where *i-Conc* use led to incorrect texts, dissatisfaction, or sometimes false perceptions of
success. However, negative effects from i-Conc consultation did not necessarily mean that i-Conc provided wrong answers or answers that were not helpful. To a large extent, the participants’ language proficiency played a role throughout the whole process of consultation ranging from formulating an effective query term that would return what was sought for, noticing what they were looking for from query results, to appropriately applying them to their own writing. However, the participants’ data demonstrated what can be called pitfalls or traps in reference resource consultation as a whole, which was also a major contributor to the unsuccessful resolutions of certain problems. Often the participants were not aware of these pitfalls, but even when they knew how to avoid them, they did not bother to do so under time pressure. The pitfalls that the participants commonly encountered can be broadly attributed to a “lack of rigor” (Kennedy & Miceli, 2001, p. 77) in evaluating query results, lack of flexibility in trying out different options, and limitations of the resources themselves. Each of these is discussed below.

4.4.2.1 Lack of rigor in evaluating query results

Unlike in classroom settings where a teacher would be able to monitor students’ tool use and care would be taken to provide enough time for students to go through each step of the consultation cycle, the participants in this study worked independently on one of their authentic writing assignments, with some participants working under severe time pressure, and thus often could not afford to closely observe query results. This lack of rigor in observation was manifested broadly in the following two ways.

First, although frequency served as a handy criterion for evaluating query results in many cases as discussed above, the participants sometimes fell into the trap of simply choosing the most frequent item from the results without examining its possible meaning differences from less frequent alternatives or whether it was actually used in the context that they intended. This happened often when the participants used COCA and JTW, where a query typically returns a list of linguistic items that meet the conditions of the query with frequency information. In these resources, one can check the sentences in which these items are actually used by clicking each item on the list. Without clicking into these actual concordance lines, the participants often just selected the item at the top of the list and applied it to their writing, which sometimes turned out to have a different meaning from what they had intended at the outset. It is this overreliance
on frequency information that contributed greatly to false perceptions of successful problem solving by the participants. For instance, in HP38, Jinho intended to extract from *i-Conc* a preposition that can be used with *concentrate* when it means “to be present in large numbers”. However, he chose *on* simply because it was presented as the most frequent preposition that occurs with the verb *concentrate* without realizing that *concentrate on* is not what he intended to mean.

Second, as opposed to serendipitous discoveries discussed above, in many cases, the participants did not notice what they were not looking for—i.e., they were only focused on (certain aspects of) the target item and neglected to pay attention to its morphosyntactic details or the patterns of its co-texts. This often resulted in ungrammatical text formulation even when the participants chose the right item for the context because of, for example, missing an article or plural marker. For instance, in GP33, Goeun rightly chose the verb *pose* for a verbal collocate of the noun *challenge*. However, even though most of the concordance lines she perused had the pattern of *pose + determiner + challenge(s)* but she applied it as *posed challenge* in her writing.

4.4.2.2 Lack of flexibility in trying out different options

The participants showed a tendency, in varying degrees, to develop their own routines in *i-Conc* consultation. In many cases, routinized consultation worked as an efficient heuristic in solving certain types of problems. However, by using only specific ways of representing problems, and using only certain resources and options, the participants missed out on opportunities to reach better solutions or the same solutions more effectively.

Some participants, especially Yumee, and Shia, showed a strong preference for verification queries over elicitation ones. They tended to perform verification queries even for problems for which elicitation queries would have been more helpful. Verification queries, which are intended to test a hypothesis about a match between L2 form and meaning, are in general easier to formulate than elicitation queries as the writer just types in a query term she wants to verify without having to use wildcards or special query syntax as is often required in formulating an elicitation query. For some query purposes such as Coll or uAlt, however, a verification query does not get the writer very far when its results do not verify her hypothesis. She then reaches a dead end and to continue the problem solving she needs to go back to the early steps of the consultation cycle and run extra queries. An elicitation query, by contrast, if
strategically formulated, returns different alternatives to choose from and often more appropriate alternatives than the item the writer initially had in mind. An example can be drawn from YP47, where Yumee carried out a verification query of *made its identity* on Google to see if the verb *make* can be a good collocate of the noun *identity*. The query returned a considerable number of matches and she accepted it with confidence. However, if she had carried out an elicitation query for the same purpose using COCA or JTW, she would have realized that there are many more typical verb collocates of *identity* such as *establish, create, develop, or forge* and chosen one of them. As such, by overusing verification queries, the participants did not tap into the full potential of the resources that can offer multiple, and often more pertinent, alternatives to choose from.

In the same vein, some participants preferred to use only the resources that required no complicated query syntax and return results quickly. In doing so, they used less effective resources for given query purposes, or consulted *i-Conc* for only a limited set of purposes that are suitable for those preferred resources. Yumee, for example, used Google with great frequency particularly for SC purposes. In fact, her rate of positive effects by *i-Conc* was one of the highest among the participants, suggesting that her preference for specific resources and query purposes did not negatively affect her problem solving. However, as noted above, verification queries, especially SC queries more often than not led to additional queries. Indeed, her number of queries per problem space was 3.03, the largest of all participants, meaning that Yumee had to perform more queries to reach a solution within a problem space than did any other participants. If she had been more flexible in choosing more varied resources for a wide range of purposes, her problem spaces would have been shorter and arrived at possibly better solutions.

4.4.2.3 Limitations of resources

In addition to the participants’ lack of rigor and flexibility in *i-Conc* use, the limitations of the resources featured in *i-Conc* themselves contributed to a certain extent to unsuccessful and unsatisfactory consultation results. As this point is discussed in detail in 4.5, only the limitations of major resources are touched upon below.
First, COCA offers a variety of options and query operators for sophisticated searches. However, some participants had problems in navigating through different options and formulating appropriate query syntax, which often made them abandon the consultations.

Google, as noted a few times already, searches the Web, which contains English language data enormous both in amount and variety, often returns a great number of matches even for non-standard language items. This quality means that even if the writer enters a non-standard or ungrammatically formulated query term, Google returns a relatively large number of its instances, giving false confirmation on the correctness or appropriateness of the target item. In GP38, Goeun performs an SC query on Google to see if the word sequence *will I be ever capable* was acceptable. Google returned many matches of the string and Goeun got reassurance that she could use it as it was. However, if she had other resources such as COCA or even GS, she would have found few or no instances of it as the word order is non-standard (as opposed to *will I ever be capable*).

By contrast, GS and CSE, included in *i-Conc* to mitigate the weaknesses of Google Web, often returned few or even no instances of even relatively common items due to their small corpus size. Some participants took this as evidence of non-use of the target item, and other participants ceased to consult these resources altogether after a few attempts.

### 4.5 Participants’ evaluations of *i-Conc*

In this section, I present how the participants evaluated *i-Conc* and its component resources in their interviews and final survey. The evaluation is broadly reported on two dimensions: (a) the benefits and challenges the participants experienced while using *i-Conc* as writing assistance and (b) the changes they perceived while using the reference suite in their confidence, attitudes, and approaches with respect to English academic writing. As in other sections above, the findings presented here are group perceptions commonly expressed by more than one participant. Individual participants’ evaluations and perceptions of *i-Conc* are provided in case studies in Chapter 5.

#### 4.5.1 Benefits and challenges of using *i-Conc*
In the final interview and survey, the participants were on the whole positive in their evaluations of i-Conc as writing assistance, saying that it helped them to solve specific types of problems that arose while completing their self-chosen authentic academic papers. However, they also provided a fairly long list of the challenges and difficulties, though not shared by all, they experienced with i-Conc as a whole and each component resource. In what follows, major benefits and challenges are discussed first, and then specific rankings the participants gave to the individual resources in terms of usefulness and difficulty of use are presented.

To begin with the benefits, when asked what they considered the biggest advantage of using i-Conc, all participants gave similar answers along the lines of easy access to a variety of mutually complementary language reference resources in a single interface. Most of them said that they found it especially useful to have multiple resources from which they could get corroboration of the accuracy/appropriacy of given linguistic items and that if they had had to visit each individual resource site and keep open multiple tabs or windows of web browsers, they would not have used all of the resources they consulted during the study. Another overall advantage is that given the resources to consult, some participants (Goeun and Yumee) have become more willing to go beyond their current lexical and syntactic repertoires for English writing and try new ways of expressing their ideas.

As for the benefits or advantages of each component resource, first, Bilingual was received as an all-purpose reference resource with its versatility of functioning as a dictionary, thesaurus and even corpus. In addition, the speed and ease with which it can be consulted was a major contributor to its frequent use among the participants. COCA was considered a concordancer that provides considerable amounts of examples that are reliable in terms of their sources, and the sophisticated search options and operators allowed the user to elicit and verify phraseologies and grammar in creative ways. JTW was especially seen as a simple-to-use but powerful tool for finding and checking collocations. As for Google, most participants agreed that it was a quick and easy tool in checking the presence and frequency of a word sequence of any length. The final resource that drew a common response was GS. Although they used it much less frequently than the above-discussed resources, the participants all mentioned that it was a resource to consult to check the register of given linguistic items. The rest of the resources, though preferred for different reasons by individual participants, did not draw common responses about their advantages.
With regard to challenges in using i-Conc, the participants’ responses were more complex and concrete compared to their responses about its advantages. Asked about the single greatest challenge of using i-Conc as a whole, most participants pointed out the time-consuming nature of its consultation processes. The participants associated its time-consuming nature with two broad causes. The first cause stemmed from the nature of Web-based resources. Some concordancers run by non-commercial institutions, particularly JTW, were much slower in returning query results (at least during the study period) than were online resources running on powerful commercial search engines (such as Google and Naver); moreover, they crashed occasionally due to their relatively unstable systems coupled with an overall unstable Internet connection on the university-wide wireless network. Being used to high-speed Internet and commercial search engines, almost all participants expressed frustration over sometimes having to wait more than 10 seconds (in the case of JTW) for the results to be loaded. The second, more complex and fundamental, cause came from the nature of reference resource consultation, especially concordancing. As in most previous studies on corpus-tool use, some participants were frustrated by the time taken to go through the consultation cycle, especially when they consulted the concordancers. That is, they sometimes found it too time consuming to choose a resource and a search option, figure out the right query syntax, and go through concordance lines to pick out what they were looking for. This second challenge is also closely related to another disadvantage of i-Conc identified by some participants. Going through the consultation cycle, particularly query results evaluation, added to their cognitive load, which was already heavy while they attended to the various aspects of the writing task at hand. The connection between extra work needed for i-Conc consultation and its effect on cognitive load was illustratively articulated by Ian:

[…] it's time-consuming. That's one reason but another reason was that it gives too many concordance lines and it's too much work for me to read through them and to find contexts that match my intention. You have to judge whether a concordance line matches your intended meaning and I often found that there are as many concordance lines that match as ones that don't. So it's not easy to find concordances that exactly match what I intend to express. So because of that, it seems that I only look at forms rather than trying to detect subtle differences in nuance hidden in concordances. (Ian, Final Interview)

By extension, Jae and Ian pointed out that while engaging in solving a specific lexical or grammatical problem for sometimes more than several minutes, i-Conc consultation therefore occasionally interrupted their flow of thoughts and distracted them from the main task at hand,
which was converting much broader units of ideational content into text. This point was also poignantly expressed by Ian:

[…] what bothered me a little was that it can easily distract you from your writing. You cannot afford to be interrupted while writing down what is shaping up in your head in terms of content. But when I use this tool, I am tempted to use each of the four tools here to confirm the accuracy. I'd want to use JTW, want to use allintext on Google, etc. So at one point I decided that it may be better to keep tool consultation to a minimum while drafting and use it at the revising stage. (Ian, SR3)

Another important and related challenge was identified by half the participants (Jae, Goeun, and Ian). They shared the view that consulting i-Conc is different from receiving writing support from human tutors or experts in terms of the cognitive burden imposed on the L2 writer. Park (2010) likened interaction with a corpus system with “dialogic negotiation” through which the user zeroes in on a solution. However, these participants’ responses suggest that although there are interactive input-feedback transactions between the user and the tool, consulting i-Conc is far more cognitively taxing and time-consuming than for example dialogic negotiation with a human tutor because it is still fully the writer’s responsibility to make judgments and decisions at every step of the consultation cycle. Moreover, even at the end of intensive interaction, they are often not fully confident about the solution they found. The following remark by Jae illustrates this point well:

Interaction with the tool is not as satisfactory as interaction with actual writing consultants or editors. In the latter case, I can ask questions when I don’t understand something. But with this tool, everything is heavily dependent on you. It’s almost 100% your responsibility to find an answer to the problem. (Jae, Final Interview)

Besides what has been discussed above, the participants mentioned that there was some redundancy in the resources, some of which overlap in functions, and that i-Conc did not offer much in supporting discourse-level problem solving.

As for COCA, all participants but Yumee said that it was the most difficult resource to use in i-Conc as it required special query syntax and option selection. Some participants (Shia and Ian) found its corpus neither large nor specialized enough and its interface very unfriendly to the user. As for JTW, every participant pointed out its slow loading time and frequent crashes as the biggest challenge in its use. The participants were also unanimous in their evaluation of
GS and CSE. They considered that these specialized Google search engines run on too small corpora (compared to Google Web Search) to return sufficient and satisfactory query results.

In summary, the participants’ overall evaluation of i-Conc was to a large extent positive. They agreed that it helped them solve certain types of problems they encountered while writing in English. They particularly liked the accessibility and mutual complementarity of the different resources provided in a single space. However, i-Conc was not without difficulties in its use. The participants overall found it challenging to deal with extra time and the cognitive load i-Conc consultation involved.

Given what the participants considered the pros and cons of the i-Conc resources above, it is worthwhile to consider how the participants actually rated the i-Conc resources in terms of usefulness and difficulty of use, and how that correlated with the actual frequency of resource use. In the final survey, the participants rated i-Conc resources for usefulness and difficulty of use on a scale of 1 to 5. Table 4.11 presents the averages of these ratings and juxtaposes them with the actual use of each resource.

Table 4.11

<table>
<thead>
<tr>
<th>Resource</th>
<th>Average rating</th>
<th>Resource</th>
<th>Average rating</th>
<th>Resource</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilingual</td>
<td>4.00</td>
<td>Bilingual</td>
<td>1.33</td>
<td>Bilingual</td>
<td>25.6</td>
</tr>
<tr>
<td>COCA</td>
<td>3.67</td>
<td>Monolingual</td>
<td>1.50</td>
<td>COCA</td>
<td>25.4</td>
</tr>
<tr>
<td>Google</td>
<td>3.50</td>
<td>GS</td>
<td>1.67</td>
<td>Google</td>
<td>18.1</td>
</tr>
<tr>
<td>JTW</td>
<td>3.50</td>
<td>Google</td>
<td>1.67</td>
<td>JTW</td>
<td>11.1</td>
</tr>
<tr>
<td>Monolingual</td>
<td>3.50</td>
<td>Thesaurus</td>
<td>1.83</td>
<td>Monolingual</td>
<td>7.8</td>
</tr>
<tr>
<td>GS</td>
<td>3.33</td>
<td>CSE</td>
<td>2.17</td>
<td>Thesaurus</td>
<td>5.2</td>
</tr>
<tr>
<td>Thesaurus</td>
<td>2.83</td>
<td>JTW</td>
<td>2.50</td>
<td>CSE</td>
<td>3.5</td>
</tr>
<tr>
<td>CSE</td>
<td>2.33</td>
<td>COCA</td>
<td>3.83</td>
<td>GS</td>
<td>3.3</td>
</tr>
</tbody>
</table>

The extra cognitive burden that query syntax formulation and query results interpretation often placed on the participants, as described above, is well reflected in Table 4.11. As can be seen in the middle column, COCA, the most sophisticated concordancer in i-Conc, was considered the most difficult resource to use. However, as the first column shows, COCA, despite its difficulty of use, was still considered one of the most useful resources, and was frequently consulted in actual use (see the last column of Table 4.11). The frequency of use to a
large extent corresponds to the perceived usefulness as can be seen in the almost identical order of resources in the two columns.

4.5.2 Changes in confidence, attitudes, and approaches

Asked if i-Conc use helped them have greater confidence in their English academic writing, all participants responded positively. Jae’s remark on confidence was particularly illustrative:

[i-Conc] helps me have confidence a lot. In the past I felt helpless, but now I think I can produce, to some degree, readable and refined texts as long as I have time and motivation to use the tool. I would say i-Conc is a suite that has almost everything that is available online which helps users edit their writing on their own. (Jae, Final Interview)

The positive responses by the participants were along the lines that i-Conc use provided them with the multi-source corroborations of the lexical, grammatical, and stylistic accuracy and appropriacy of what they intend to say and thereby reassurance that they could use it with confidence.

However, some participants made clear that using the reference suite did not bring noticeable changes in their overall confidence in English academic writing as a whole because they still “feel burdened” when they had to write in English (Goeun, and Shia) and struggled with discourse-level problems (Ian).

The participants said that i-Conc use did not bring major changes in their attitudes and approaches to English academic writing either, suggesting that using a reference tool for relatively short periods of time would not have a major impact on their writing practices, which had been developed over years. However, although the participants did not explicitly identify them as changes, there were a few important changes that were commonly found in the participants’ responses. First, all participants but Jinho made remarks along the lines that i-Conc encouraged them to be more adventurous in their language use and go beyond safe expressions they always relied on and to try out new vocabulary and syntactic structures. This observation was well expressed by Goeun and Shia:
[i-Conc] allows me to go beyond my current linguistic repertoire, which is quite limited and underdeveloped, to try new ways of expressing my thoughts. As I have studied English for many years, now I tend to seek only safe ways of saying things. I feel like my English is becoming fossilized. With this tool, you can check whether your experimental expressions can actually be used or not. (Goeun, Final Interview)

[While using i-Conc] I tried a bit more to pursue variety in my vocabulary. I made efforts to use words that I didn't usually use. (Shia, Final Interview)

Second, some participants said that while using i-Conc they became more attentive to lexical choice and grammar while drafting a paper than they were before (Yumee and Shia), and while doing so, they also became more strategic and sophisticated in their reference resource use (Yumee, Goeun, Jinho and Ian) by using the tools for purposes for which they had never consulted resources before. Third, using concordancing in particular made some participants (Jae, and Yumee) aware of and attentive to the lexico-grammatical aspects (e.g., collocation and colligation) of vocabulary they used, which in turn increased their uses of the concordancers. Finally, some participants (Jae and Shia) mentioned that they felt they had greater independence as L2 writers, and would eventually reduce their dependency on outside support such as help from writing centres. On the changes she experienced, Shia said:

I tried to take this as an opportunity to wean myself from the writing centre, and as it happened I couldn't use their service for this paper. So I reviewed my paper more closely than usual. In the past, I just trusted tutors at the writing centre to catch what I missed. This time I wrote my paper pretending I won't get any help from them. So I took more care about accuracy. (Shia, Final Interview)

Overall, these changes, although seldom explicitly identified by the participants as such, can be seen as the positive effects of i-Conc on the participants’ confidence and approaches regarding English academic writing. This tendency can be confirmed in the participants’ responses to a question about whether they would continue to use the reference suite as writing assistance. They all answered that they would do so, if not with all resources featured in it, while writing, especially, important papers for which a great deal of revisions and refinements are required. Table 4.12 summarizes these group findings presented above. Unique but important findings and insights from individual participants are presented in the next chapter.
Table 4.12
Summary of Participants’ Evaluations of i-Conc

<table>
<thead>
<tr>
<th>Resources</th>
<th>Advantages</th>
<th>Challenges</th>
</tr>
</thead>
</table>
| i-Conc    | - Access to mutually complementary resources in a single space (all participants)  
- Greater willingness to venture beyond the current linguistic repertoire (Yumee, Goeun)  
- Reassurance about accuracy, and appropriacy of the vocabulary (Jae, Yumee, Goeun, Shia, Ian)  
- More attentive to accuracy and appropriacy (Jae, Yumee)  
- Better awareness and noticing of problems (Jae, Yumee)  
- Increased sense of autonomy in L2 writing (Jae, Shia)  
- Improved search skills (Yumee, Goeun, Jinho)  | - Time-consuming (Jae, Shia, Ian)  
- Cognitively taxing/distracting (Jae, Goeun, Ian)  
- Uncertainty about the solutions found (Jae, Ian)  
- Lack of interactivity (Jae, Goeun)  
- Not helpful beyond lexical and grammatical aspects of writing (Goeun, Ian)  
- Overlap in functionalities among resources (Jinho, Ian)  |
| Bilingual | - Multi-purpose resource (Yumee, Goeun)  
- Fast and easy to use (Jae, Yumee)  
- Searches on English-Korean equivalents (Jae, Shia, Ian)  | |
| COCA     | - Sophisticated searches (Jae, Yumee, Jinho)  
- Examples from reliable sources (Jae, Jinho)  | - Difficult to use due to query syntax and operators (Jae, Geun, Shia, Ian)  
- Not specialized enough for searches on discipline-specific terms (Shia, Ian)  |
| JTW      | - Simple to use (Jae, Goeun)  
- Organized display of results (Jae, Goeun, Ian)  | - Slow loading and frequent crashes (Yumee, Jinho, Shia, Ian)  |
| Google   | - Quick and easy resource for verification (Yumee, Goeun, Shia)  | - Too much informal, non-standard data (Jae, Goeun, Ian)  |
| GS       | - Resource for checking academic register (Jae, Goeun, Ian)  | - Too small in corpus size (Jae, Yumee, Goeun, Ian)  |
| CSE      | | - Too small in corpus size (Jae, Yumee, Goeun, Ian)  |
CHAPTER FIVE

CASE STUDIES

The previous chapter presented group findings, addressing research questions to provide a general account of how the participants used i-Conc while completing an authentic writing assignment in English. However, this general picture has been drawn by averaging and abstracting the results from individual participants, each of whom showed wide variability from one another in many aspects of their i-Conc use. In this chapter, the research questions are answered from each individual participant’s perspective in the form of case studies. Using the general accounts provided in the previous chapter as a reference point, the six case studies illustrate in depth where individual participants’ i-Conc use and its evaluation diverged from and converged with one another, and thereby identify potential factors that contributed to individual differences and similarities. Each case study is structured in three parts: a brief introduction about the participant, a comprehensive table that shows major quantitative results from the participant’s i-Conc use, followed by explanation of these uses, and then the participant’s evaluation of i-Conc as writing assistance and their perceived changes in various aspects of their English academic writing brought about by their uses of i-Conc.

5.1 Case study 1: Jae

5.1.1 Background

At the start of the present study, Jae was the oldest participant (37) and was at the latest stage of degree completion, working on his PhD dissertation in the field of human resources and adult education. Despite plenty of writing experience during his graduate studies, however, his confidence in English academic writing was the lowest among the participants; he felt he had difficulty across all areas, especially with academic style and register (see Table 5.1). In the initial interview Jae’s response to the question about the biggest source of frustration with respect to English academic writing illustrated clearly the attitudes he had towards writing in English: constant doubt about whether what he has just written down on paper makes sense, and
further whether his texts are at the level expected of a PhD student. In other words, Jae said he was frustrated by his inability to self-evaluate words he choose to use, what a sentence structure he formulated would sound like, or whether a paragraph he just finished would be seen as coherent by the target audience. He summarized this as inability to write an academic paper in English on his own. To compensate for this perceived lack in writing ability, Jae sought help from a variety of writing support programs run by his university such as writing centres and peer tutors throughout his PhD studies. Jae said he made it a rule to have all major papers and proposals reviewed by native English speakers, and when he could not do so it was because of time conflicts, for which he experienced severe frustration. Through years of using those services, Jae had developed strategies for effectively getting the particular types of help he needed from writing consultants and tutors, some of which were presented earlier in the previous chapter.

5.1.2  i-Conc use

Table 5.1

An Overview of Jae’s i-Conc Use

<table>
<thead>
<tr>
<th>Overall confidence in L2 academic writing/ Perceived difficulty in different areas of writing (on a scale of 1(lowest) to 5 (highest))</th>
<th>Confidence</th>
<th>Vocabulary/Usage</th>
<th>Syntax/Sentence Construction</th>
<th>Academic style/ Register</th>
<th>Rhetoric/ Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problems and Queries</th>
<th>No. of problems and queries</th>
<th>Linguistic categories&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Query purposes&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Primary query purposes&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Resources consulted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem space (PS)</td>
<td>54</td>
<td>lex 52.0%</td>
<td>Mn 31.6%</td>
<td>Coll 29.6%</td>
<td>COCA 41.8%</td>
</tr>
<tr>
<td>Query (Q)</td>
<td>98</td>
<td>le-gr 40.8%</td>
<td>Coll 26.5%</td>
<td>uAlt 20.4%</td>
<td>Bilingual 28.6%</td>
</tr>
<tr>
<td>Q/PS</td>
<td>1.81</td>
<td>gram 7.1%</td>
<td>uAlt 15.3%</td>
<td>Mn 20.4%</td>
<td>Google 8.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>styl 12.2%</td>
<td>Reg 11.2%</td>
<td>Equi 14.8%</td>
<td>CSE 5.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Equi 10.2%</td>
<td>GU 7.4%</td>
<td>Monoling. 4.1%</td>
</tr>
<tr>
<td>Types of problems and queries&lt;sup&gt;d&lt;/sup&gt;</td>
<td>No. of abandoned problem spaces</td>
<td>SC 8.2%</td>
<td>SC 7.4%</td>
<td>Thesaurus 4.1%</td>
<td></td>
</tr>
<tr>
<td>Conf</td>
<td>57.4%</td>
<td>SC 8.2%</td>
<td>GU 6.1%</td>
<td>AP 5.6%</td>
<td>GS 3.1%</td>
</tr>
<tr>
<td>Comp</td>
<td>42.6%</td>
<td>7 (13%)</td>
<td>AP 5.1%</td>
<td>Reg 1.9%</td>
<td>JTW 1.0%</td>
</tr>
<tr>
<td>V</td>
<td>56.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>48.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem solving results</th>
<th>positive effects</th>
<th>negative effects</th>
<th>satisfied consultation</th>
<th>dissatisfied consultation</th>
<th>successful resolution</th>
<th>unsuccessful resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>from i-Conc use</td>
<td>75.9%</td>
<td>11.2%</td>
<td>77.8%</td>
<td>22.2%</td>
<td>88.9%</td>
<td>11.2%</td>
</tr>
</tbody>
</table>

Top 3 resources (in order of preference)
Given his background, it may not be surprising that Jae was one of the most enthusiastic participants about the idea of using i-Conc as a supporting tool for problem solving during English composition. Although he was not the most frequent user of the suite among the participants in terms of the number of queries carried out, Jae took up the largest number of problems (54) to solve using i-Conc during the three screen recordings. His i-Conc consultations were effective overall and helped him solve problems successfully, evidenced by his high rate of positive effects from using i-Conc. A number of patterns characterize his i-Conc consultation. Firstly, Jae was the most frequent user of COCA in terms of percentage among the six participants. Before participating in the study, he had only used Naver, the same bilingual dictionary (Bilingual) featured in i-Conc as a reference resource. However, while participating in the study, Jae developed a preference for COCA, which despite its complex interface he considered a very useful tool having features and options that Bilingual, his hitherto main reference resource, did not provide. Jae specifically cited as the advantages of COCA multiple alternatives to choose from, frequency information, and plenty of authentic examples extracted from reliable sources.

Secondly, even though Jae had some notions of formulaic language and lexico-grammatical co-occurrence before the study, it was in the tutoring sessions that he was introduced to the concepts of collocation and colligation and over time he realized he had a tendency to translate Korean expressions into English word by word. Reflecting his concerns about whether his writing is not academic enough and, in his own words, “gives away his identity as a non-native speaker from Asia,” Jae carried out a great number of queries especially to look for words (or word combinations) that would sound natural and more academic, evidenced by the relatively high proportions of Coll, uAlt and Reg in terms of his query purposes.

Thirdly, Jae carried out a relatively very small number of SC queries given that SC was the most frequent purpose the participants on the whole consulted i-Conc for (see 4.2.2.2 for group findings on query purposes). This pattern may also reflect Jae’s low confidence in the
texts he formulated and his cautious approach. Rather than directly testing his hypotheses using SC queries, Jae preferred to see what the reference resources had to suggest using elicitation queries. This cautious tendency was also evident in his verification queries, which were mostly done for the purposes of checking the match between an L2 form and intended meaning (Mn), a purpose which stands in contrast to the patterns shown by other participants, whose verification queries included a much higher proportion of SC queries.

Fourthly, closely related to the third pattern discussed above, Jae used elicitation queries very strategically. As presented in the previous chapter, along with Ian, Jae frequently used the strategy of starting Conf problem spaces with an elicitation query and thereby saving himself extra queries. Jae also often formulated creative query syntax to solve unique problems (for examples, see JP66, and JP35 presented in Table 4.9 in Chapter 4).

Lastly, however, Jae’s i-Conc consultations did not always lead to successful resolutions of problems. As presented as one of the pitfalls in i-Conc use, Jae sometimes relied too much on frequency information, using it as the only criterion for choosing an item to apply to his writing without checking whether the item delivered the intended meaning. Another major cause for unsuccessful problem solving was that Jae sometimes forgot the query operators and syntax required on COCA due to a rather long hiatus in i-Conc use between his screen recordings and so he had to relearn them.

5.1.3 Evaluation of i-Conc

In evaluating i-Conc as writing assistance, Jae provided insights into the most varied aspects of the reference suite among the participants. To begin with what he considered the advantages of i-Conc use and positive changes i-Conc use brought to him, first, like other participants, Jae pointed to the easy accessibility to mutually complementary resources from a single interface, which makes it possible to corroborate query results from one resource with another. Secondly, using i-Conc, especially the concordancers, Jae said, helped him raise his lexico-grammatical awareness that a text is not merely a concatenation of single words that are complete in meaning and, by extension, Jae increased his awareness of differences between

---

9 For personal reasons, Jae could not steadily use i-Conc while working on the paper he had chosen for the study during the participation period.
Korean and English in the ways of expressing the same ideational content. This is well expressed in his following remarks:

I also realized that there are typical and appropriate expressions and sentences in a given context. Connecting individual words that have meanings I intend is not enough, that's what beginning level language learners do. (Final Interview)

I think I translate what has been conceptualized in Korean into English without being aware of me doing it. (Informal interview after SR1)

Thirdly, Jae suggested that consulting the reference resources while writing not only meets the writer’s immediate needs by helping solve the problem at hand but also provides opportunities for language learning through repeated exposure to L2 forms and contexts in which they occur.

All these advantages and positive changes, Jae explained, combined to boost his confidence to write in English “on his own” and become an independent writer. He expressed this sense as follows:

I realized as well that this tool makes me more autonomous, and independent in solving problems while writing although time-consuming and rather unidirectional. (Final survey)

However, as implied in the quotation immediately above, Jae also experienced challenges and difficulties using i-Conc. First, as for i-Conc as a whole, Jae said it was a challenge to him to stay motivated to use the suite throughout the entire duration of writing, particularly at the drafting stage. The biggest reason was, he said, that as with other participants, the extra time and cognitive exertion required to perform queries. However, this time-consuming and cognitively taxing nature of i-Conc seemed to have bothered Jae more than others. In his endeavor to publish and present in as many venues as possible, Jae saw every paper he worked on as a draft to be developed into a conference or journal paper through extensive revision with the help of native speaker editors. Jae thus suggested that pursuing precise phrasing at the draft stage can be somewhat a waste of time and that i-Conc could therefore be better used at the revising and editing stages of writing. The second and closely related reason he gave for losing motivation was that using i-Conc during drafting sometimes
interrupted the flow of thoughts negatively, affecting his content development of the section or paragraph he was working on. Jae described these challenges thusly:

It was hard to stay motivated to use it throughout the whole process of writing. Unless you need to produce a really accurate text like when you have an impending deadline for your conference paper or course paper, you wouldn't need to use it. Using Naver (bilingual dictionary) alone you can produce "accurate enough" papers. (Final Interview)

[…] and it tends to distract your attention from the content you're developing. It can interrupt the flow of your thoughts. […] converting your ideas into text and having them written down before they fly away is the most important thing while drafting. So I don't think I can use it for long and steadily while drafting. (Final interview)

Jae also raised a question about the nature of interaction between the writer and the reference resources. As someone who was so used to getting help from native speaker tutors and writing centre consultants, Jae saw his interaction with i-Conc as rather unidirectional in that it was still solely his responsibility to formulate the right questions and pick out the right answers from query results.

Jae shared the limitations he found from individual resources within i-Conc as well, which were in large part congruent with those pointed out by other participants. However, as the most frequent user of COCA, Jae pointed out a few important limitations of COCA that no other participants did. First, although there are benefits in accessing ample examples of how a given item used in a sentence, sometimes COCA returns too many concordance lines that are displayed randomly rather than displaying the most typical or representative contexts first. Second, there are no distinctions in concordance lines returned by COCA between the different senses of a polysemous word. Jae realized that frequency information provided on COCA can thus be misleading sometimes.

Cognizant of these limitations, Jae suggested that despite many useful features and benefits of COCA, it should be complemented by more traditional resources such as bilingual dictionaries. Indeed, Jae chose Bilingual and COCA as the top two resources he would use most frequently in the future:

I like COCA and NAVER best. I think I use them with similar frequency. Both have most clear strengths. Naver allows Korean queries and is very fast. COCA gives you usage examples. So they are two opposing poles in terms of strengths and weaknesses. I tend not to use resources that fall between the two poles. (Final interview)
This observation shows that Jae was well aware that one type of references is not necessarily better than others, but they can be used most effectively in combination as several scholars have argued (Flowerdew, 2009; Frankenberg-Garcia, 2005).

Jae’s evaluation of the limitations of i-Conc as a problem solving aid for English academic writing were at times so strongly expressed that they seem to contradict his overall positive evaluation of the suite. Particularly, Jae expressed repeatedly his doubts about its utility as a supporting tool while drafting a paper. However, toward the end of the final interview, Jae shared how different it felt to revise the sections he had written with and without i-Conc. The sections he drafted using i-Conc were much clearer in what they intended to convey than the sections drafted without the suite. In saying this, Jae suggested that consulting reference resources during text formulation may not be a waste of time but rather draws the writer’s attention more keenly to form-meaning mapping, and thereby shapes more clearly the idea the writer tries to convey. Asked about whether he would continue to use i-Conc in the future, Jae responded as follows:

If you go to great lengths to make each sentence accurate, refining them as you go, later you might end up regretting doing so when you have a clearer idea about what each paragraph should be about, and how they should be organized once you finish drafting, because that would normally mean you have to rewrite those sentences later. So all the work you did for each sentence would get wasted. That's how I felt when I started using i-Conc. However, now I believe you can learn the language from using the tool. So even if sentences are rewritten later, you still learn from the process. So I am going to try to use it [while drafting]. (Final interview)

Jae realized that the extra time and cognitive load involved in using the reference resources while drafting would eventually pay off as he came to pay greater attention to words he uses in his papers and therefore learned more about the target language and to develop ideas more articulately in the process.

5.2 Case study 2: Yumee

5.2.1 Background

When Yumee volunteered for the study, she was in the fourth semester of her PhD program in educational technology, already with considerable academic writing experience,
having written a dozen of major course papers and several conference papers. Yumee was the only one among the participants that explicitly stated that she enjoyed writing academic papers in English. Although Yumee’s self-rating of her confidence in English academic writing was not particularly high, she said that she did not find the four aspects of writing any more difficult than did other participants (see Table 5.2).

Asked about whether she set specific goals when writing an academic paper, Yumee responded that she tried to draft papers early enough to send them to her peers and academic advisor for feedback. However, during her first few semesters in the program, Yumee realized that she had been writing in English just as she would in Korean. That is, her English sentences had been convoluted, connected with multiple relative clauses and heavy with all the ornate vocabulary, which often received the negative feedback that they were unintelligible. With the beginning of the semester, Yumee made it an overall goal for her English writing to make her sentences simple and short and to use only vocabulary she was comfortable using for the sake of clarity. As for surface-level editing, Yumee mainly focused on coherent presentation of ideas and did not try to get every vocabulary and grammar right in her papers, which she believed is extremely difficult to achieve as an L2 writer and slows down writing as a whole. Like Jae, Yumee considered that course papers are drafts that will eventually be developed into conference and journal papers, and so expending too much time and energy on the surface-level accuracy of a draft is not very “cost-effective.”

As hinted at above, Yumee received writing support from the most varied sources among the six participants. Unlike other participants, she was very open to seeking advice and feedback on her papers from her peers and academic advisor. Particularly her academic advisor was a major source of writing support, which was not the case with other participants. The advisor helped Yumee with surface-level editing and proofreading for all her major course papers, even the ones that were not for her courses. This may have been part of the reason that Yumee was not that concerned about the surface-level accuracy of her English writing even though she found Vocabulary/Usage more difficult than other aspects (see Table 5.2 below).
Yumee consulted an online bilingual dictionary frequently while writing, even before participating in the study, and her i-Conc use during the study period proved to be by far the greatest in terms of the number of queries, accounting for more than half of all queries carried out by all participants (1,739 out of 3,170 queries). Even when considering only the screen-recorded queries, she performed the largest number of queries or 121.

An overarching theme, which Yumee referred to as “efficiency over perfection”, runs through her i-Conc consultation patterns. What she meant by “efficiency” in reference resource consultation was to take time-effective consultation paths that bring “good enough” solutions that involve the least time and distraction. Yumee’s efficiency-oriented i-Conc use was well demonstrated by the consultation patterns she developed over the period of the research. First, although her problems were relatively evenly divided between confirmatory and compensatory
types, Yumee carried out many more verification than elicitation queries: at a ratio of about 7 to 3. This tendency can be explained by her preference for testing her intuitive hypotheses about target L2 forms through a series of SC queries—even within compensatory problem spaces—over exploring alternatives to choose from. As can be seen in Table 5.2, SC accounted for almost 50% of all query purposes for which she consulted i-Conc. As discussed in Chapter 4, formulating elicitation queries and evaluating their results tend to require extra time and cognitive exertion from the writer. For the sake of efficiency, Yumee strategically avoided these types of queries and did a series of SC queries instead. Adding to this, what she referred to as her personality, as described in her following remarks, contributed to this consultation pattern:

> I think personality matters. I just trust my intuition and easily get convinced and happily accept the results without much reflection. (Final Interview)

This inclination is well evidenced by the high satisfaction rate of her i-Conc consultations (87.5%), which was the highest of all participants, and the low rate of abandoned consultations (10%).

Second, Yumee’s efficiency-oriented tendency was also manifest in her choice of resources. As Table 5.2 shows, Yumee used mainly three resources, which combined made up more than 95% of all her queries. As an educational technology major who had considerable experience developing and using online educational software programs, she gave up using the resources that were slow and unstable (JTW), returned few or no instances of queried items (GS, and CSE), or created extra cognitive demands from extra L1-L2 translation (Monolingual, and Thesaurus) after several attempts to use them. She did not find particular functions from these resources that were so unique and useful as to compensate for their shortcomings. Indeed, as she continued to use the suite, she developed a routine in consulting i-Conc. Yumee described her routine and approach to reference resource consultation as follows:

> I have developed my own routines or search patterns. I usually start with Naver (Bilingual) and get an idea there and to be sure, run another query on Google. If greater sophistication is needed, then I go to COCA. There are overlapping functions between resources. So I rarely go to JTW, or the thesaurus. If you invest more time and effort,

---

10 Yumee said that she has a tendency to mentally translate into Korean the definitions and examples given for the query term in these resources.
you can get better results but I am not so keen to do that. I am satisfied with my routines.

(Final Interview)

Yumee’s i-Conc consultation patterns can be summarized as using in highly routinized ways the resources that she judged were fast, and returned sufficient results, and using them mainly for testing her intuitive hypotheses (i.e., SC). Yumee believed this approach gave her the greatest efficiency in consulting reference resources.

However, Yumee’s preference for efficiency over precision was not without disadvantages. Her routinized consultation patterns proved to be effective overall in problem solving as confirmed by the high rate of positive effects from her i-Conc consultations (75%). However, her particular routines may not be the most efficient ways of consulting i-Conc. As discussed in the previous chapter, verification queries can often bring the writer to a dead end or necessitate further queries. Indeed, as can be seen in Table 5.2, Yumee carried out an average of 3.03 queries per problem space, the largest number among the participants. Her pursuit of efficiency ironically resulted in performing more queries within her problem spaces than did other participants. A large number of queries within a problem space do not necessarily mean inefficiency, however. As also discussed in the previous chapter, Yumee tended to habitually overuse verification, especially SC queries, using only a limited set of resources, even when elicitation queries on other resources would have been more efficient and helpful (e.g., YP47, see its description in Section 4.4.2) by saving her extra queries and offering potentially more appropriate alternatives to choose from.

Occasionally, her tendency to “easily accept” the query results led to a lack of rigor in evaluating the results and modifying the retrieved items to fit the intended context in her writing. Yumee fell into a trap of concluding her query hastily by checking only the first two or three concordance lines. Even when she retrieved the right item from the query results, Yumee sometimes ended up formulating a grammatically incorrect text by applying the item without modifying its aspect and tense to fit her intended context. This pattern can explain her relatively high rate of the problems she mistakenly considered successfully resolved (12.5%).

In summary, Yumee developed highly routinized i-Conc consultation patterns that were more aimed to find “good enough” solutions quickly than to go to great lengths to get precise solutions. This efficiency oriented consultation paid off with her high rate of successful resolution of problems. However, satisfied with this routinized i-Conc consultation patterns,
Yumee did not explore ways of using the suite even more efficiently and reaching better solutions.

5.2.3 Evaluation of i-Conc

Yumee was the most positive of all participants in evaluating i-Conc, being the only one who gave a rating of 5, the highest point, to i-Conc for its usefulness for English academic writing. Likening it to “a gift box of assorted goodies,” she said the biggest advantage of i-Conc was that it provides easy access to multiple resources from a single interface without having to open multiple windows while writing. What she considered an incidental advantage to this quality was that the presence of multiple resources in the suite often made her want to use more than one resource for confirmation. This behaviour helped her, if not drastically, improve the accuracy of her texts and thereby gain confidence in her writing. By way of example, she mentioned a recent paper of hers she had just gotten back from her academic advisor with suggestions for surface-level editing. She said that there were many fewer comments on surface-level editing in the sections she had written using i-Conc than in those written without the suite.

Asked if there were any changes in her attitudes and approach to English academic writing while using the suite, Yumee responded that there were two major changes. First, before the study she had only used an online bilingual dictionary and occasionally Google while writing mainly for simple purposes of finding L2 equivalents and confirming the acceptability of certain linguistic items. But while using the suite, she realized that reference resources can be used for more sophisticated purposes such as finding collocates, which she had never done before, and she even learned to use the resources that she had been using in more sophisticated ways such as enclosing search terms with double quotes on Google. Second, Yumee became more attentive to the accuracy and appropriacy of words and phrases she used in her writing by checking them against the resources. Particularly she used to borrow terms and phrases uncritically from their sources, but she now got into the habit of checking them using i-Conc to see if they were appropriate to the contexts intended in her writing. Yumee said:

If it can be seen as a change in terms of approach, now I tend to check what I borrow from books, and authors for accuracy and appropriacy, which I didn't used to do. I just accepted them without much reflection. (Final Interview)
Yumee did not say much in terms of disadvantages about i-Conc as a whole, reflecting her highly positive attitude to it. However, she pointed out the limitations of the resources she chose not to use. As briefly discussed above, Yumee saw no unique benefits for using JTW, Monolingual, Thesaurus, GS, and CSE.

Instead, Yumee considered Bilingual, Google, and COCA to be sufficient for solving most types of problems that arose during her written assignment. Among these three her preference for Bilingual was particularly strong as she considered it the most efficient resource of all in terms of speed and relevance of information provided. Yumee took full advantage of the functions and features provided in the resource, using search strategies that were not used by other participants. For example, using sentence translation pairs (English and Korean) provided in the resource as a parallel corpus, she carried out SC queries. She said about her favorite resources:

Using Naver, Google and sometimes COCA, brings the greatest cost-effectiveness for me. I care most about efficiency. Naver is not just a bilingual dictionary, but it has all the functionalities that JTW and Thesaurus have. It provides synonyms, antonyms, etc. Naver is more like an all-purpose resource. (Final Interviews)

5.3 Case Study 3: Jinho

5.3.1 Background

Jinho, the youngest participant, had quite a different background from other participants. While the five other participants all came to North America to study in their current graduate programs, Jinho had immigrated to Canada with his family when he was 14 years old. He finished his high school and undergraduate education in Canada before entering his current MA program in information studies. Despite the relatively long residence in Canada, however, he considered Korean his primary language in important domains of his life such as home and church.

When it comes to academic writing in English, despite much writing experience during high school and college, Jinho had been struggling since starting his MA program to deal with the much higher standards required for graduate-level course papers, which became a major
source of stress and frustration. In his first semester in the program, two of his course papers failed to get a “pass” grade because, according to the course instructors, they were lacking in both content and presentation. Since then he had been making various efforts to improve his English academic writing, taking non-credit academic writing courses, and using writing centre services. Participating in the present study was part of his endeavor to enhance his abilities to edit his own papers. As for specific areas of English writing to improve, he answered “grammar and sentence construction,” which was also somewhat different from other participants, who were more concerned about vocabulary or discourse-level aspects of writing.

5.3.2 i-Conc use

Table 5.3

An Overview of Jinho’s i-Conc Use

<table>
<thead>
<tr>
<th>Overall confidence in L2 academic writing/ Perceived difficulty in different areas of writing</th>
<th>Confidence</th>
<th>Vocab/Usage</th>
<th>Syntax/Sentence Construction</th>
<th>Academic style/ Register</th>
<th>Rhetoric/ Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>(on a scale of 1(lowest) to 5 (highest))</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problems and Queries</th>
<th>No. of problems and queries</th>
<th>Linguistic categories</th>
<th>Query purposes</th>
<th>Primary query purposes</th>
<th>Resources consulted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem space (PS)</td>
<td>46</td>
<td>lex 63.4%</td>
<td>Mn 32.3%</td>
<td>Mn 34.8%</td>
<td>COCA 36.6%</td>
</tr>
<tr>
<td>Query (Q)</td>
<td>93</td>
<td>le-gr 10.8%</td>
<td>SC 20.4%</td>
<td>Reg 15.2%</td>
<td>Monoling. 22.6%</td>
</tr>
<tr>
<td>Q/PS</td>
<td>2.02</td>
<td>gram 3.2%</td>
<td>uAlt 12.9%</td>
<td>SC 8.7%</td>
<td>Bilingual 10.8%</td>
</tr>
<tr>
<td>Types of problems and queries</td>
<td>No. of abandoned problem spaces</td>
<td>Conf 73.9%</td>
<td>Equi 6.5%</td>
<td>Coll 6.5%</td>
<td>Thesaurus 2.2%</td>
</tr>
<tr>
<td>Conf</td>
<td>26.1%</td>
<td>10 (21.7%)</td>
<td>sAlt 4.3%</td>
<td>NE 4.3%</td>
<td>GS 1.1%</td>
</tr>
<tr>
<td>V</td>
<td>71.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>29.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem solving results</th>
<th>positive effects from i-Conc use</th>
<th>negative effects from i-Conc use</th>
<th>satisfied consultation</th>
<th>dissatisfied consultation</th>
<th>successful resolution</th>
<th>unsuccessful resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>63.0%</td>
<td>15.2%</td>
<td>58.7%</td>
<td>41.3%</td>
<td>78.2%</td>
<td>17.4%</td>
</tr>
</tbody>
</table>

Top 3 resources (in order of preference)

COCA – Monolingual – CSE

Note. a. lex: lexical, le-gr: lexico-grammatical, gram: grammatical, styl: stylistic
b. Coll: collocation, Equi: L2 equivalent, Mn: intended meaning, Oths: others,
   NE: noun ending, sAlt: simple alternative, SC: simple confirmation, uAlt: upgrading alternative.

See Table 3.8 for detailed description of these codes.
Table 5.2 shows some salient patterns of Jinho’s i-Conc use that contrasted with other participants. His problems were predominantly of the Conf type, which accounted for almost 75% of his total problems. The queries he carried out were also of the verification type more than 70% of the time, suggesting that he consulted i-Conc to “check” much more frequently than to “explore”. When it comes to primary query purposes, Jinho ran a relatively small number of queries for Coll, Equi, and SC purposes, which were the most frequent queries for the group. Instead, he performed Mn queries most frequently, and his proportion of Mn queries was much higher than those of other participants. In addition, though not presented in Table 5.3, more than 50% of his queries were single-word queries, which was again a very high proportion compared to most other participants. These patterns—i.e., the high proportion of single-word, Mn verification queries—suggest that Jinho tended to consult the resources in i-Conc including the concordancers for the purposes that the more traditional resources (e.g., dictionaries) are typically consulted for. Looking at the distribution of his queries across the linguistic categories, Table 5.3 shows that Jinho was predominantly concerned with lexical problems (almost 65%) while running only a small number of queries of a lexico-grammatical nature (less than 11%), once again suggesting that he used i-Conc more for dictionary functions than concordancer functions.

Turning back to his consultation patterns for the dictionary component of i-Conc, he showed another major difference from the general trend observed in other participants. He preferred Monolingual to Bilingual, which was the most frequently consulted resource in the group average. His proportion of Bilingual consultations was the lowest among the participants. Likewise, he did not carry out as many Equi queries as other participants. His explanation of the small proportion of Equi queries and Bilingual use was that he had held this belief, influenced by his parents and what was taught back in Korea, that to improve English proficiency, one should immerse himself in English and reduce the use of the mother tongue to a minimum. With this belief, Jinho had been using most of the time monolingual dictionaries and formed the habit over his lengthy residence in Canada of checking different English alternatives to convey a given idea using a monolingual dictionary rather than conceiving it in Korean first and then finding English equivalents in a bilingual dictionary.

Turning to how his i-Conc use bore on his problem solving performance, as shown in Table 5.3, the rate of positive effects by i-Conc use was relatively low at 63.0% (compared to
the 70.8%, average of the other participants), and the rate of the queries he was satisfied with was even lower at 58.7% (cf. 75.9%, the average among other participants). In addition, Jinho was shown to abandon a relatively high percentage (21.7%) of consultations midway.

Screen recording observation and Jinho’s recollections revealed that there were three major contributors to these patterns. First, in a number of cases, Jinho started off a problem space without clearly defining the given problem first—i.e., defining exactly what to look for through *i*-Conc consultation. Instead, he merely entered a single word on different resources in sequence with no specific anticipation for the results in mind. For example, in HP22, Jinho queried the noun *gatekeeper* on JTW, Monolingual, and Thesaurus in sequence seemingly with no clear purposes and no logical links between each query such as a subsequent query building on the previous one, which was one of the major query strategies used by other participants discussed in the previous chapter.

Second, his overall lack of understanding about the unique functions of the resources resulted in a considerable number of instances of mismatch between his query intentions and the resource or option he chose. Despite the introduction of corpus linguistic concepts and basic instructions in concordancing provided at the beginning of the study, Jinho had an incomplete understanding of how and for what purposes the new type of reference resources (i.e., concordancers) can be used. In other words, he did not clearly grasp what kinds of problems can be effectively and uniquely addressed with each of the concordancers. For example, his responses in the final survey revealed that he saw as the main advantages of COCA its abilities to present an ample amount of sentence examples and to show where a specific word comes in a sentence, which are not usually considered the primary advantages of using a concordancer. This caused certain difficulties for him in formulating query strategies, that is, choosing the most effective resource or option for solving a specific problem. In a number of his concordancer consultations, he chose a resource or option in a seemingly haphazard way, rather than choosing the one that is usually consulted for the given problem type. For example, in HP40, Jinho queried the noun *betterment* on JTW, and later recalled that he had done so to see if it was a frequently used word even though JTW is a resource mainly intended to get collocations rather than frequency information about single words.

Third, Jinho’s difficulty using query syntax/operators and selecting the right search options in COCA was not unique, but it was particularly pronounced with him. He was notably
unable to become fully familiar with the query syntax/operators until the last screen recording; he formulated query terms in a syntax that either was not allowed or would not return the intended query results in over 30% of his COCA queries. One query syntax error he frequently made, for example, involved the synonym operator, i.e., [=word], which returns the synonyms of the word. He had an incomplete understanding of the functions and limitations of the operator and, in many instances, used it more as a place holder or put a word in the bracket that was too general in meaning with multiple senses or a word too specific or technical to have any synonyms. In HP53, he formulated the query recent [=one] on COCA with an intention to find alternatives of the indefinite pronoun one, which he felt sounds informal. This query returned many irrelevant results (such as recent single, and recent unique) because the concordancer also returned the synonyms of the word one in all its other senses.

As a potential underlying cause for the patterns discussed above, his relative lack of (meta-)linguistic knowledge may have made it difficult to define problems and formulate proper query terms. As the youngest participant, who learned English in ESL settings in Canada, Jinho had not been exposed to all the (meta-)linguistic concepts and terms either in English or in Korean as much as other participants had experienced during their education in Korea, and he himself considered grammar an area that he needed to improve the most. Although he seemed to have no problem understanding the linguistic content of the tutoring provided at the beginning of the study, observation of his screen recordings revealed that he often had trouble formulating query syntax for target items in a specific POS as he was often confused about the names of different word classes.

With respect to query results evaluation, he showed a tendency to uncritically accept query results and make decisions based on them too readily without further queries for corroboration in different resources. In some cases, for example, he readily took no instance in the results as evidence of the non-existence or non-use of the target item without reflecting on the specificity of the query or the limitations of the resource consulted. For example, in HP31, he queried the word re-label on Monolingual to see if it is acceptable to use the word in the context of “relabeling the existing categories.” When the resource returned no matches of the word, Jinho took it as evidence that it is not a legitimate word, and rephrased it “label again,” which can have a different meaning from re-label. He was not aware that dictionaries usually do not have derivatives as separate entries.
These patterns of Jinho's difficulties should not be equated with difficulties in English writing itself. The majority of lexical, grammatical and stylistic infelicities in his texts were minor errors that did not affect the overall intelligibility of his written communication. In large part, Jinho resolved the problems successfully (see in Table 5.3 the rate of successful resolution, 78.2%). What has been described above suggest then that certain resources such as COCA require a solid grip on (meta-)linguistic concepts for successful use.

5.3.3 Evaluation of i-Conc

Like his overall i-Conc use, Jinho’s evaluation of the suite was unique in some important aspects compared to those of other participants. First, unlike other participants, he said that the use of i-Conc helped him speed up the process of consulting reference resources and that he was thus able to spend more time on other aspects of writing such as idea development. More specifically, he said i-Conc allowed him to access various resources in a single interface and therefore saved him from the usual cumbersome task of locating and retrieving the needed information from his offline and online reference resources:

In the past, I used to consult different resources for different purposes: dictionaries for meanings, some materials I keep on computer for sentence structures. But now that I can do those things with i-Conc, the process is faster with no interruption. (Final Interview)

However, he did not mention the other side of the time issue caused by the extra cognitive load imposed by i-Conc use, which most participants pointed out as one of the major challenges of using i-Conc.

With respect to his confidence, Jinho responded that the use of i-Conc did not boost his overall confidence in academic writing but that it increased his confidence in search skills, which he believed constitute an important part of writing skills:

But it seems that confidence improved in terms of my ability to use search tools rather than my writing ability itself. Search skills are part of your overall competence. We cannot know everything, the meaning of every word, and all the details of English grammar. So I think it's a different domain of skills that helps you find what you don't know. (Final Interview)
Asked about the challenges in i-Conc use, Jinho responded that there was considerable overlap in functions among the resources, and so he often had to stop his writing to figure out where to go to perform the intended query. Indeed, when performing queries, Jinho was often observed to move the mouse cursor restlessly along the different tabs of i-Conc for up to several seconds. As demonstrated in the previous section, another major difficulty he experienced was familiarizing himself with the query syntax and operators required to use on COCA. Jinho admitted that while he saw great value in using COCA for supporting his own writing, his lack of control over COCA’s specific query syntax and options affected the overall outcome of his i-Conc use. As he put it,

I haven't gotten used to COCA's query syntax, like you have to insert parentheses, or an asterisk. If I had gotten the syntax right, I would have gotten much better results. That's what I found the most difficult and regrettable. (Final Interview)

As to future use, he said that although he was not able to use it exactly in the ways it was intended to be used, he would continue to learn more about the resources and consult them for his future papers.

5.4 Case Study 4: Goeun

5.4.1 Background

Goeun was one of the two participants who were specializing in language education and therefore were the most familiar with (meta-)linguistic concepts explained in my training, surveys, and interviews. However, her self-rated confidence in English academic writing was not any higher than other participants, and Goeun found the different areas of writing difficult as seen in Table 5.4. Despite or maybe because of her overall high proficiency in English, however, her concerns for English writing were less about the discrete aspects of writing but about the overall ability to pull together competence in these aspects to produce sophisticated texts that faithfully convey the ideas she intend to state. In this respect, Goeun considered her English “underdeveloped,” and “not sufficient” to express her ideas with all the nuances and subtle shades she intended to convey.
Unlike other participants, Goeun could rarely afford to seek writing support from any sources because she tended to finish drafting her papers at the last minute before their deadlines. Indeed, she was the only one who did not use writing centre services among the participants. Instead, she had been using a wide variety of reference resources to get immediate support while writing in English. She used most of the types of reference resources for writing before participating in the study and was the only one who had prior experience using corpus tools (COCA and Collins COBUILD Corpus).

5.4.2  

*i-Conc use*

**Table 5.4**

An Overview of Goeun’s i-Conc Use

<table>
<thead>
<tr>
<th>Overall confidence in L2 academic writing/ Perceived difficulty in different areas of writing (on a scale of 1(lowest) to 5(highest))</th>
<th>Conf</th>
<th>3</th>
<th>4</th>
<th>Syntax/Sentence Construction</th>
<th>3</th>
<th>4</th>
<th>Academic style/ Register</th>
<th>4</th>
<th>Rhetoric/ Organization</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of problems and queries</td>
<td>Linguistic categories&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Query purposes&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Primary query purposes&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Resources consulted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem space (PS)</td>
<td>25</td>
<td>lex</td>
<td>38.5%</td>
<td>Equi</td>
<td>26.9%</td>
<td>Equi</td>
<td>40.0%</td>
<td>Bilingual</td>
<td>40.4%</td>
<td></td>
</tr>
<tr>
<td>Query (Q)</td>
<td>52</td>
<td>le-gr</td>
<td>50.0%</td>
<td>SC</td>
<td>15.4%</td>
<td>Coll</td>
<td>16.0%</td>
<td>Google</td>
<td>21.2%</td>
<td></td>
</tr>
<tr>
<td>Q/PS</td>
<td>2.08</td>
<td>gram</td>
<td>15.4%</td>
<td>Coll</td>
<td>13.5%</td>
<td>AP</td>
<td>12.0%</td>
<td>JTW</td>
<td>17.3%</td>
<td></td>
</tr>
<tr>
<td>Types of problems and queries&lt;sup&gt;c&lt;/sup&gt;</td>
<td>No. of abandoned problem spaces</td>
<td>POS</td>
<td>9.6%</td>
<td>sAlt</td>
<td>8.0%</td>
<td>Monoling.</td>
<td>3.8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conf</td>
<td>32.0%</td>
<td>SH</td>
<td>5.8%</td>
<td>GU</td>
<td>4.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>72.0%</td>
<td>6 (24%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>42.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>61.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Problem solving results

<table>
<thead>
<tr>
<th>positive effects from i-Conc use</th>
<th>negative effects from i-Conc use</th>
<th>satisfied consultation</th>
<th>dissatisfied consultation</th>
<th>successful resolution</th>
<th>unsuccessful resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>56.0%</td>
<td>20.0%</td>
<td>64.0%</td>
<td>36.0%</td>
<td>64.0%</td>
<td>20.0%</td>
</tr>
</tbody>
</table>

Top 3 resources (in order of preference)

JTW – Bilingual – GS

*Note.*  

a. lex: lexical, le-gr: lexico-grammatical, gram: grammatical, styl: stylistic  
See Table 3.8 for detailed description of these codes.
Of all six participants, Goeun carried out the smallest number of queries, or 52, to solve the fewest problems, or 25. This can be explained by a few factors. First, she had a very tight schedule during the study period with multiple deadlines to meet for the courses she was taking. Thus, Goeun was not able to allot much time for the writing assignment she had chosen for the study and could not afford to revise and refine the paper using i-Conc as much as she had hoped to. Second, although she had been using varied types of reference resources, Goeun tended to use those reference resources only when absolutely necessary especially when drafting. Goeun explained that while drafting, she usually gets ideas for vocabulary and syntactic structures from the articles and books that she reads for a given writing assignment, and so she felt there was not much need for consulting reference resources. Third, due to her high proficiency in the target language, she may have encountered fewer problems that she judged to be soluble by consulting i-Conc than did other participants.

Despite her small numbers of problems and queries, several salient patterns were observed in Goeun’s i-Conc use. The problems she took up to solve with i-Conc were predominantly of the Comp type, making up 72% of all problems, which is the highest of all participants. Her queries were to a lesser degree dominated by elicitation queries (61.5%).

With respect to query purposes, Table 5.4 shows another unique pattern of Goeun’s i-Conc use. The most frequent purpose her primary queries were performed for was Equi, accounting for 40% of all purposes, which is far higher than the percentages of Equi in other participants’ data. This trend is in large part attributable to the specific type of writing Goeun did during her screen recordings. The writing assignment Goeun chose for the study was a paper on her personal trajectory of language learning and links between the personal experiences and a certain theory of language learning. Therefore, the bulk of the paper was of a narrative nature, with recollections of her life back in Korea, describing the various emotions she went through in the process of learning and using English, memories of which were experienced and retained in large part in her L1. Thus Equi queries aimed to find matching L2 forms comprised the largest proportion of her queries.

As for her choice of resources, Goeun consulted Bilingual the most frequently, which was not surprising given the high proportion of Equi queries. Another salient pattern is that her COCA use was almost non-existent in the screen-recorded sections of her paper. Instead, her JTW use was relatively high in terms of proportions compared to most other participants.
To look at how Goeun’s i-Conc consultation patterns bore upon her problem solving results, her rate of positive effects by i-Conc use and rate of successful problem resolution were the lowest of all participants. On the other hand, her proportion of abandoned consultations was the highest at 24%. This pattern means that Goeun abandoned i-Conc consultation for almost one in every four problems. These results were quite surprising given her high proficiency in the target language and her familiarity with all reference resources within i-Conc. However, this unexpected, relatively poor performance can also be traced back to the type of writing task she was doing and what it implied for her i-Conc use in terms of the type of problems and query purposes. To begin with, as presented in Chapter 4, problems of the Comp type turned out to be hard to solve and more likely to end up abandoned across all participants. Secondly, because of the narrative nature of her paper, many of the Comp problems she took up to tackle with i-Conc concerned L2 vocabulary and syntactic patterns to faithfully describe certain classroom scenes retrieved from her memory and accompanying emotions. But those L2 forms turned out to be elusive to extract through i-Conc consultation. Reflecting the difficulty she had performing these queries, Goeun gave up formulating the given idea into text in four out of the six abandoned problems. Thirdly, Goeun was one of the participants who had to work with a very tight deadline which did not afford her the time needed for formulating queries and evaluating results thoroughly.

5.4.3 Evaluation of i-Conc

Like other participants, she also considered the single access point to multiple reference resources as the biggest advantage of i-Conc as a whole. By making use of these multiple resources, she was able to get ideas of how to strategically solve lexico-grammatical problems that she encountered while writing in English and also to learn new search skills that she had not known before participating in the study. Another important benefit was changes in her attitude and approach to English writing. As presented already as an exemplary case in the group findings, she believed that i-Conc stimulated her to go beyond what she saw as her limited linguistic repertoire to experiment with new ways of expressing her thoughts. As a language pedagogy major, she even suggested that i-Conc would help prevent fossilization of learners’ interlanguage. As far as confidence is concerned, Goeun shared with most other participants the
view that the suite did not directly enhance overall confidence in English academic writing. However, she said that it allowed her to confirm and get reassurances about the accuracy of lexical and grammatical items she tried to use and thus helped her use vocabulary with greater confidence.

When it comes to individual resources, she found JTW particularly useful. Although she had the same difficulty as other participants caused by its slow loading and occasional crashes, Goeun liked very much its “no-frills” interface and the highly organized way in which it displays query results in clusters divided according to the syntactic relationships between the node word and collocate and the different senses of the node word. For this reason, JTW was her first choice when asked about top three resources she would continue to use after the study.

Regarding the challenges and difficulties she experienced while using i-Conc, Goeun first pointed out the differences between consulting i-Conc and consulting with human experts. She suggested that i-Conc can be effectively used only when the user has a clear idea of what she is looking for, that is, when confirming something she already knows, whereas interaction with an expert (tutor or editor) helps her find out what she needs to know through meaning negotiation. She explained:

It's quite different from consulting with actual native speakers. I can do meaning negotiation with them, getting closer to what I am looking for. But to solve problems with this tool, you have to have clear ideas about what to say and you can only confirm those ideas with the tool most of the time. (Final Interview)

This observation implies that Goeun did not consider i-Conc an effective tool for compensatory problem solving, perhaps reflecting the difficulty she experienced solving problems of the Comp type with i-Conc as discussed in the previous section.

As for individual resources within i-Conc, Goeun especially found COCA unfriendly to the user as a reference tool for a number of reasons. First, its interface was so complicated, with different parts of the screen popping up upon mouse-over, that it was often hard to navigate through the various pieces of information provided and to quickly locate what she was looking for. Second, the way COCA displays concordances seemed to her random with no distinctions between different senses of the target item. Though sharing many functionalities with JTW, COCA seemed to her more complicated for performing queries but to display the results less
effectively than JTW. For this reason, Goeun used JTW instead of COCA when she needed to consult a concordancer.

Finally, for future use, Goeun stated that she would continue to use the suite for writing, she was highly motivated to do so, but not for a course paper with a tight deadline, suggesting that as some participants noted, under time pressure of a tight deadline, it’s hard to get motivated to use i-Conc.

5.5 Case Study 5: Shia

5.5.1 Background

Shia, an MA student studying social work, was one of the two participants who volunteered to participate in the study during the summer term. As it was also her final term of the degree program, she had a demanding schedule to meet all the final degree requirements while working on the writing assignment she chose for the study, which had a much tighter deadline than those of the winter-term participants. That was the reason why her study participation period was only seven weeks, the shortest among the participants.

In the initial survey, Shia indicated the highest confidence in English academic writing of all participants, giving a 4 on a scale of 1 to 5. However, in the initial interview, she said that it was “relative confidence” that she had developed over time by comparing her papers with other students’ including those of native speaker classmates. Shia explained that as a newcomer to Canada lacking in language ability and cultural knowledge, she had to invest twice as much time and effort in writing as her native speaker classmates did. Although far from perfect in terms of grammar and vocabulary, Shia believed her papers often had more substance and were more coherent than those written by some of her native speaker peers.

Reflecting her confidence, Shia’s perceived difficulty of the different aspects of academic writing was less than other participants except for Vocabulary/Usage (see Table 5.5). She was the only one who gave the score 2 to any of the areas presented in the survey. However, there were some specific grammar items she had difficulty with. Shia said she had long been struggling with articles and prepositions in her English academic writing.
When it comes to writing support, like Goeun, Shia could hardly afford to do serious editing or revising as she tended to finish drafting a paper close to the due date. However, she instead made it a rule to take a draft of every major paper to the writing centre for feedback before submitting it. Indeed, along with Jae, Shia was one of the two participants who strategically used the writing centre services within the strict restriction on their frequency of use, by, for example, booking sessions with only the tutors who knew her writing style and weaknesses and going into a session with specific questions to ask from or about her paper.

5.5.2 i-Conc use

Table 5.5

An Overview of Shia’s i-Conc Use

<table>
<thead>
<tr>
<th>Overall confidence in L2 academic writing/ Perceived difficulty in different areas of writing (on a scale of 1(lowest) to 5 (highest))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Problems and Queries

<table>
<thead>
<tr>
<th>No. of problems and queries</th>
<th>Linguistic categories&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Query purposes&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Primary query purposes&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Resources consulted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem space (PS) 28</td>
<td>lex 26.2%</td>
<td>SC 38.9%</td>
<td>SC 46.4%</td>
<td>COCA 23.6%</td>
</tr>
<tr>
<td>Query (Q) 72</td>
<td>le-gr 59.7%</td>
<td>Mn 18.1%</td>
<td>Reg 21.4%</td>
<td>Bilingual 22.2%</td>
</tr>
<tr>
<td>Q/PS 2.57</td>
<td>gram 9.7%</td>
<td>Reg 18.1%</td>
<td>Coll 10.7%</td>
<td>Google 19.4%</td>
</tr>
<tr>
<td></td>
<td>styl 22.2%</td>
<td>Coll 12.5%</td>
<td>Equi 10.7%</td>
<td>GS 12.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Equi 6.9%</td>
<td>Mn 10.7%</td>
<td>JTW 6.9%</td>
</tr>
<tr>
<td>Types of problems and queries&lt;sup&gt;c&lt;/sup&gt;</td>
<td>No. of abandoned problem spaces</td>
<td>uAlt 6.9%</td>
<td>Art 3.6%</td>
<td>Thesaurus 5.6%</td>
</tr>
<tr>
<td>Conf 82.1%</td>
<td>4 (14.3%)</td>
<td>AP 2.8%</td>
<td>NE 3.6%</td>
<td>CSE 2.8%</td>
</tr>
<tr>
<td>Comp 17.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V 79.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E 20.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Problem solving results

<table>
<thead>
<tr>
<th>positive effects from i-Conc use</th>
<th>negative effects from i-Conc use</th>
<th>satisfied consultation</th>
<th>dissatisfied consultation</th>
<th>successful resolution</th>
<th>unsuccessful resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>71.4%</td>
<td>14.2%</td>
<td>78.5%</td>
<td>21.5%</td>
<td>78.5%</td>
<td>17.9%</td>
</tr>
</tbody>
</table>

Top 3 resources (in order of preference)

Bilingual – Google – Thesaurus

Note. a. lex: lexical, le-gr: lexico-grammatical, gram: grammatical, styl: stylistic

See Table 3.8 for detailed description of these codes.
Shia took up a relatively small number of problems, or 28, to solve with i-Conc, performing a total of 72 queries during her three screen recordings. One major pattern that stands out from her i-Conc use is the markedly high rates of confirmatory problems, 82.1%, and verification queries, 79.2%, which are by far the highest of all participants. As can be seen in Figure 5.5, this pattern is also well illustrated by the distribution of query purposes. Shia’s top three query purposes, SC, Mn and Reg, which account for more than 75%, were purposes that are typically met through verification queries. Among these three, SC was especially predominant, making up 46.4% of all her primary query purposes, meaning that almost half of her problems were initiated for SC purposes. Her Conf- and verification-oriented consultation can also be observed in her Coll queries. Unlike the general trend shown by other participants, Shia’s Coll queries were predominantly verification queries.

When it comes to evaluation of query results, one salient pattern was that in many of her SC queries, Shia took even a small number of instances of the queried word sequence as evidence of its acceptability. In other words, she showed a tendency to take the mere presence of the queried item in the query results rather than its frequency as a criterion for deciding whether to apply it to her writing.

These tendencies of her i-Conc use may be attributed in large part to the time pressure she had been under while completing the assignment. In her stimulated recalls and final interview, Shia repeatedly emphasized that she was struggling with her tight schedule and thus the “absolute lack of time” for “polishing” her sentences. To a large extent, this circumstance contributed to the pattern: she mostly used the suite to confirm the target items she had already formulated on screen or in her mind, that is, for SC purposes rather than exploring (better) alternatives that the reference resources would suggest, which would have taken a longer time and greater cognitive effort. In doing so, Shia not only consulted search engines (Google and GS) a lot, over 30% with Google and GS combined, which are typically consulted for SC, but also used COCA as a search engine performing many simple multi-word queries without specific syntax and operators. Her pattern of evaluating the query results described above—i.e., only checking the presence rather than the frequency of target items in the results—was also something of a compromise necessitated by a lack of time. Shia explained:

Under time pressure, I just get satisfied with only a few matches. Although it doesn’t tell you whether it is frequently used but shows that there are people who said/wrote it in the
way I want to. I just settle for confirming that my expression [word sequence] is not totally off. (Informal interview following the 2nd stimulated recall)

This reference resource consultation pattern driven by the lack of time seems to have worked out well for Shia’s problem-solving performance. In 71.4% of the problems, her i-Conc consultation had a positive effect, slightly over the average of all participants. However, screen recording observation revealed that Shia’s SC-dominant consultation pattern often made her fall into the pitfalls specific to SC queries. First, although Shia had little room for choice under the severe time pressure, she could have avoided some of her unsuccessful consultations by trying out elicitation queries. Some of her SC queries returned no or few matches because the queries themselves were ungrammatical (e.g., SP18 and SP29). Second, as discussed in Section 4.4.2.3, reference resources running on a huge corpus such as Google—due to the nature of the Web it searches—often gave a false confirmation of the accuracy or acceptability of a linguistic item that is non-standard, ungrammatical or grammatically acceptable but semantically incorrect by returning a considerable number of instances of them in the results.

Although not specific to SC queries, another pitfall Shia was prone to was lack of rigor in evaluating query results. Again under severe time pressure, she tended to simply check the presence or frequency of target items but stop short of confirming whether the items found in the results were used in the contexts she intended. In SP26, for example, Shia queried relapse occur on COCA and GS to see if the verb occur can be used with the noun relapse. She confirmed there were instances of that word combination and applied it to her writing. However, she did not notice from the query results that the noun relapse was mostly used in relation with diseases or medical conditions while she meant to use it in the context of discussing the recurrence of an undesirable policy with no intention of likening it to a disease.

In summary, Shia used i-Conc effectively given the severe lack of time, focusing on confirming her hypotheses about L2 forms. However, her confirmatory- and SC query-dominant i-Conc use pattern made her prone to some specific pitfalls such as simply checking the instances of the searched-for items and uncritically accepting them as evidence of their accuracy and appropriacy in the given contexts.

5.5.3 Evaluation of i-Conc
Shia selected the abundant authentic examples as the biggest advantage of using i-Conc for her English writing. But overall she did not think that i-Conc provided drastically different functions from those provided by existing reference resources. From her perspective, i-Conc was an extension of what the traditional resources offer but with more sophisticated options.

Asked if there were any positive changes she experienced in her attitude and approaches to English writing, Shia responded that although she did not use i-Conc often or long enough to feel any major changes, at least for the writing assignment she had chosen for the study, she tried to go beyond the words and phrases she had been comfortable using and explore alternative ways of expressing her ideas. As documented already in Chapter 4, Shia also explored the possibility of using the suite as a supporting tool that helps her to become a more independent writer, recognizing that there will be no more writing centre service to turn to after the completion of her university program.

On the other hand, Shia also had specific difficulties and challenges while using i-Conc. First of all, compared to the stand-alone bilingual dictionary software she had used before participating in the study, i-Conc was somewhat complicated to use with all its query options and also time-consuming because some resources took long to load their results. Given her writing process, constricted by the constant lack of time, Shia said she wouldn’t use the suite much at the drafting stage, echoing what some other participants stated in their interviews.

As far as individual resources are concerned, Shia pointed out that COCA had many search options that she would normally not use because they require special query syntax, making it too unwieldy to be used as a reference resource. She explained:

COCA has many functions and options but I didn't get to use most of them as they require you to use parentheses and quotes. It's like buying an expensive VCR which has lots of bells and whistles but all you end up using are the play and stop buttons. I wish it was more intuitive. I don't think that it's been developed in user-centered ways. (Final interview)

In addition, Shia found that COCA and JTW may be better for searching for general words and phrases than for specialized usages specific to a certain field of study or genre. Those two concordancers often returned no instances of the word sequences that she believed to be used frequently in her field. On the other hand, CSE, which was included in i-Conc for searches for domain-specific usages, ran on so small a corpus that Shia often found no instances of the target items she searched for. In the final survey, as hinted at above, she picked Google as her
most preferred concordancer because it is fast and searches the whole Web to return a great number of hits for word sequences specific to her field.

In concluding her evaluation of the suite, Shia said that i-Conc may be an invaluable tool for writers who are highly motivated or required to produce refined texts and that she would be willing to learn more about it and use it at work when she has to be a truly independent writer.

5.6 Case Study 6: Ian

5.6.1 Background

As a doctoral student in applied linguistics, Ian was one of the two most advanced L2 writers, along with Goeun, of the six participants. However, his confidence and enjoyment in English writing was not high, and he rated all areas of writing as equally difficult (see Table 5.6). Ian was frustrated by his inability to manipulate English vocabulary and syntactic structures to fit his communication needs, which he saw was a prerequisite for producing clear and effective academic texts. He said the vocabulary and syntactic structures he first chooses while composing in English seem too simple or awkward, and he thus tends to avoid using those first things that come to mind and instead to try to find alternatives.

As for goals for academic writing, Ian responded that he usually did not set a goal specific to each writing assignment, but he tried overall to write a logical, clear, engaging, and insightful paper within its deadline. Like other participants, lexical and grammatical accuracy was one of his main concerns in English writing but he never made it a goal to produce error-free texts, which he saw is unachievable for L2 writers.

To improve his overall academic skills, Ian took non-credit academic writing courses provided for graduate students as well as took advantage of one-on-one conferences with the writing instructors for discussing his own writing. Like most other participants, Ian also used the writing centre for his major course papers and scholarship applications. However, he did not find the support from them satisfactory as the help he usually received was “template-like” feedback that could be easily found in academic writing textbooks.

Ian had already been using a wide variety of reference resources for English writing, including monolingual and bilingual dictionaries, a collocation dictionary, a thesaurus, and
academic writing style books. However, he had no prior experience of using corpus tools as a reference resource.

5.6.2  

*i-Conc use*

Table 5.6

*An Overview of Ian’s *i-Conc* Use*

<table>
<thead>
<tr>
<th>Overall confidence in L2 academic writing/ Perceived difficulty in different areas of writing</th>
<th>Confidence</th>
<th>Vocabulary/Usage</th>
<th>Syntax/Sentence Construction</th>
<th>Academic style/ Register</th>
<th>Rhetoric/ Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Problems and Queries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of problems and queries</td>
<td>Linguistic categories a</td>
<td>Query purposes b</td>
<td>Primary query purposes b</td>
<td>Resources consulted</td>
<td></td>
</tr>
<tr>
<td>Problem space (PS)</td>
<td>52</td>
<td>lex 45.6%</td>
<td>Coll 43.0%</td>
<td>Coll 44.2%</td>
<td>JTW 35.4%</td>
</tr>
<tr>
<td>Query (Q)</td>
<td>79</td>
<td>le-gr 49.4%</td>
<td>Mn 13.9%</td>
<td>sAlt 15.4%</td>
<td>COCA 20.3%</td>
</tr>
<tr>
<td>Q/PS</td>
<td>1.52</td>
<td>gram 12.7%</td>
<td>Equi 11.4%</td>
<td>Equi 15.4%</td>
<td>Bilingual 15.2%</td>
</tr>
<tr>
<td>Types of problems and queries c</td>
<td>No. of abandoned problem spaces</td>
<td>SC 6.3%</td>
<td>Mn 5.8%</td>
<td>Thesaurus 6.3%</td>
<td></td>
</tr>
<tr>
<td>Conf</td>
<td>46.2%</td>
<td>Art 5.1%</td>
<td>SC 5.8%</td>
<td>GS 1.3%</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>53.8%</td>
<td>13 (25.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>40.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>64.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Effects on writing**

<table>
<thead>
<tr>
<th>positive effects from <em>i-Conc use</em></th>
<th>negative effects from <em>i-Conc use</em></th>
<th>satisfied consultation</th>
<th>dissatisfied consultation</th>
<th>successful resolution</th>
<th>unsuccessful resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>69.2%</td>
<td>5.7%</td>
<td>69.2%</td>
<td>30.8%</td>
<td>92.3%</td>
<td>5.7%</td>
</tr>
</tbody>
</table>

**Top 3 resources** (in order of preference)

JTW – Monolingual – COCA

*Note.* a. lex: lexical, le-gr: lexico-grammatical, gram: grammatical, styl: stylistic
b. Art: article/determiner, AP: argument pattern, Coll: collocation, Equi: L2 equivalent,
Mn: intended meaning, sAlt: simple alternative, SC: simple confirmation, uAlt: upgrading alternative.
See Table 3.8 for detailed description of these codes.

Ian took up 52 problems to address with *i-Conc*, a relatively large number compared to the group average of about 40. The number of queries he performed for those problems were 79, however, meaning that he carried out an average of 1.52 queries per problem space, the fewest among the participants. In fact, observation of his screen recordings reveals that like Goeun, Ian did not consult *i-Conc* as often as most other participants. The reason that the number of problems he took up was higher than the group average is that he screen-recorded his writing for
much longer than did other participants (see Table 3.7 in Section 3.3.2.5). With regards to problem and query types, Ian showed a noted preference for compensatory and elicitation types respectively, which resembles the pattern shown by Goeun.

Coll took up the lion’s share of the purposes Ian consulted i-Conc for, accounting for more than 40% of all query purposes in terms of both total and primary queries. Following Coll in the primary query purposes were sAlt, Equi, and uAlt, which are also the purposes that are usually met through elicitation queries. In contrast, SC, one of the most common query purposes in other participants’ data made up a relatively small proportion in his i-Conc use.

Reflecting the predominance of Coll queries, JTW proved to be the top resource Ian consulted, as it is a concordancer specialized in retrieving collocational relationships a given word has with other words. Indeed, Ian consulted JTW the most frequently of all participants. Another noticeable pattern is that he consulted Google search engines less than other participants.

To look at the effects of i-Conc use on writing, Ian’s rate of positive effects by i-Conc use was at about the group average, showing again (along with Goeun) that high writing proficiency does not necessarily lead to a high rate of successful i-Conc consultations. Showing another similarity with Goeun, Ian abandoned i-Conc consultation in a relatively large proportion of his problems or in about one in every four problems and instead sought solutions on his own. However, his overall success rate of problem solving was very high at 92.3%, meaning that he successfully resolved most of these problems on his own.

The predominance of Coll queries and high rate of abandoned consultations in Ian’s i-Conc use data may be in part explained by a tendency he developed over the course of using the suite: a tendency to avoid distraction and reduce cognitive load. As will be discussed in the next section, Ian found consulting concordancers in i-Conc cognitively taxing as it required him to strategically formulate query terms, read through concordance lines, and evaluate the query results. Thus when it came to concordancing, Ian focused on what he believed it can be most efficiently and effectively used for, i.e., Coll, sAlt and uAlt. For example, a Coll query on JTW and COCA returns a list of word combinations first rather than directly presenting concordances containing the word combinations. Ian often picked out what he judged to be the most appropriate choice from the lists without clicking into each combination to see its concordance.
A close look at his abandoned consultations reveals that most of those problems were initiated to find alternatives. As discussed in Chapter 4, in a large portion of the problem spaces initiated for sAlt and uAlt purposes, the participants tended to give up i-Conc consultation midway when they were not satisfied with the query results. They already had an L2 form to express the intended meaning and they just fell back on it when they were not able to find better-sounding alternatives from i-Conc. As briefly noted in the previous section, Ian had a tendency to cast into doubt the first English word or phrase that came to his mind when expressing a concept, and then to try to find alternatives that sound better. This led to a relatively high number of sAlt and uAlt queries, which in turn led to a high number of abandoned consultations.

5.6.3 Evaluation of i-Conc

Ian stated that i-Conc was a collection of the best linguistic reference resources freely available on the Web and that allowing the user to access them in a single interface was its biggest advantage. However, as he had been using a great variety of reference resources even before participating in the study, Ian said, there was no dramatic change over the period of the research in the ways he consulted the reference resources or the purposes he consulted them for. Nonetheless, Ian found JTW and COCA particularly useful for collocation searches as they provided far more collocational information than did the collocation dictionary he used to consult. He said that while the collocation dictionary, obviously due to space restriction, has only major words especially nouns and verbs as node words and shows the combinations these words form, JTW and COCA can have a word in any class as the node. At the same time, between the two concordancers, Ian strongly preferred JTW as it does not require any special query syntax and displays the results in a very organized way. He also found Monolingual very useful for checking the specific grammar characteristics of a given word.

Asked about the biggest disadvantage or challenge in using i-Conc, Ian echoed other participants’ view that consulting i-Conc can be distracting and interrupt the flow of thoughts, particularly while drafting a paper, when the writer’s cognitive resources are focused on capturing (often fleeting) ideas into text. Formulating a query and perusing concordance lines to figure out whether the target item is used in the context he intended were sometimes time
consuming and physically and emotionally “draining” as it was not always easy to do so, and there was no guarantee that he would eventually reach a solution if he persisted. When Ian was first introduced to the corpus tools in i-Con during the tutoring sessions, he was excited by their potential for supporting his writing. However, over time he realized that a simple query on them does not offer a solution on a silver plate, but that he needed to invest extra time and effort to familiarize himself with the resources and develop search and evaluation strategies on his own, which sometimes overwhelmed him. Ian’s following remark illustrates this point:

In the beginning when I saw how corpora work, I was amazed with excitement that I can access real language, but when I actually tried to use it for my writing I didn't get clear ideas how to do it. (Final interview)

This observation suggests that even highly advanced L2 writers like Ian may need more help with what O’Sullivan (2007) termed “corpus consultation literacy.” Indeed, Ian stated that the tutoring sessions provided at the beginning of the study were insufficient to meet his needs. He expressed his wishes for a more ongoing and hands-on guidance for i-Con use:

I wish there were somebody right beside me to pinpoint my mistakes. Rather than leaving all the decisions, judgements to the user, a teacher or tutor can give feedback at every step of using the tool. It's frustrating that there is no one who can help me figure out the best ways to use the tool for specific cases. (Final interview)

One specific area in which Ian found the resources as a whole lacking was their ability to offer functional alternatives. Given that one of his major concerns while engaging in English writing was finding better alternatives to what he initially had in mind or a better way to express the same idea, he wanted to elicit various forms to serve the same function (e.g., acknowledging a limitation) from the resources that can range from a single word item, and a phrase, to even a sentence. However, i-Con offers only semantic alternatives in the same word class or linguistic unit as the original item.

When it comes to individual resources, Ian pointed out that the COCA corpus was not specialized enough to be used for a specific domain and its interface is too clunky to use as a reference resource with all its complicated options. Comparing it with JTW, which he thought is designed for language learners, Ian said COCA can be better used by linguists. As for CSE, he pointed out that despite its good intentions to help graduate students to perform queries more
specific to their own fields of study, it often returned no instances of familiar phraseologies frequently used in his field. He said:

CSE is the one that is intended to be a specialized corpus but its size is too small hardly giving you enough hits. When we read articles, there are ah-ha moments when we come across great expressions that we want to use in our own writing later. I hope corpus tools can provide those kinds of expressions. (Final interview)

Asked if there were any changes in his approaches and process regarding his English academic writing, Ian mentioned no major changes in those aspects. However, he said that i-Conc provided him with “a wider spectrum of choice” for the language he can use in his writing. He explained:

I seem to try different things that are made available by various resources in i-Conc. The Google operator "allintext" is a good example. I never used that before. In other words, it gives me a wider spectrum of choice I can choose from, rather than changing my writing process, practices, or style. (Final interview)

As for impact on his confidence in English writing, Ian responded that i-Conc gave him reassurance that he can solve more efficiently many of the vocabulary and usage problems he encounters while writing, so that he now can focus more on other aspects of writing such as coherence and organization. However, he said that i-Conc did not give any major boost to his overall confidence as he had the belief that however sophisticated a supporting tool is, the extent to which the use of that tool—or technology in general—for writing can support a writer is in large part determined by the writer’s control of it and language proficiency. He said:

Basically I believe that technology cannot solve all the problems and however sophisticated the technology gets, it is after all the human that reads, and writes. (Final interview)

Concluding his evaluation of i-Conc, Ian said that he still sees great potential for i-Conc especially the corpus tools for helping intermediate-proficiency language learners to write on a general topic.
CHAPTER SIX

CROSS-CASE ANALYSIS: INTERVENING FACTORS INFLUENCING i-CONC USE

The present chapter addresses the fifth question guiding this research: What are the factors that contribute to individual differences, if any, in the participants’ interactions with the suite? This analysis builds on the accounts in Chapters 4 and 5, respectively, of group and individual uses of i-Conc in terms of types of problems and queries, resources consulted, problem solving performance, the extent to which the participants were satisfied with their i-Conc consultations, and their evaluations of i-Conc as writing assistance. The present analysis, however, assumed that certain findings presented in the two previous chapters may have arisen from, or interrelated with, complex interactions between multiple, intervening factors beyond group or individual behaviors. Moreover, there may be differences in the extent to which each of these factors may have affected each individual participant, so even those participants who showed similar patterns along many variables may have interacted with i-Conc for different reasons. For example, the particular patterns shown by each participant in their types of problems and queries (see Figure 6.1) cannot easily be traced back to one common factor but rather appear to be affected by multiple factors with differing degrees of influence, some of which extend beyond the immediate environment of i-Conc. Lastly, rather than dichotomies of two extreme categories into which the participants can be classified (e.g., either impulsive or reflective, see below), I wish to understand these factors as continua, along which participants’ behaviors may be placed depending on differing tendencies, situational factors, and interactions among them. To this end, I have synthesized themes from the survey questionnaire and interview data into factors corresponding to three fundamental dimensions of writing: writer, text, and context (cf. Cumming, 1990, 1998).
Figure 6.1. Percentages of problem and query types in each participant’s i-Conc use.

6.1 Writer factors

The first writer factor evident in the participants’ data is the participants’ **overall goals and expectations** for their own academic writing. Applying two contrasting goal orientations often discussed in educational psychology to L2 writing development, Cumming (2012) made a distinction between *mastery*-oriented students, who aspire to “craft a fully self-satisfying composition, and to learn new things and extend their own abilities in the process of writing” and *performance*-oriented students, who “wish to finish their assignments in a satisfactory way […] then get on to other things” (p. 137). Asked about their goals and expectations in the initial survey and interview, all participants responded that while they usually do not start a writing assignment with an explicit goal or expectations, they all wish to produce a clear and coherent text that delivers the intended ideas effectively. However, their ideas about how to achieve that overall goal in terms of language use were somewhat different from one another. For example, particularly mindful of the gap in the extent to which they can deliver their ideas in their L1 and L2, Goeun and Ian were more attentive than other participants to how to express their ideational content in English in the right tone with the appropriate nuance. Yumee, in contrast, accepted the gap and strategically put her priority on communicating her ideas and intended meanings clearly by using “simple” and “good enough” language rather than “embellished vocabulary and
convoluted sentences”, such as would characterize her L1 prose. Yumee considered this to be the most time-effective way of completing an English writing assignment.

To draw a parallel with Cumming’s distinction above, Goeun and Ian may be considered to have more of a mastery goal for their writing assignment whereas Yumee’s orientation is more analogous to performance goals. These differing goals and expectations of the participants seem to have affected the types and purposes of their queries and also problem solving performance. Goeun and Ian’s i-Conc uses were characterized by high proportions of compensatory problems and elicitation queries, and also of abandoned consultations, all of which illustrate their struggles to search for accurate, appropriate, or better L2 forms for their intended meanings. In contrast, Yumee’s i-Conc use was dominated by long verification query sequences, in which she repeatedly tested her hypotheses about L2 forms for the given meanings.

Another factor contributing to these general patterns may be linked to the learner’s cognitive or learning styles. Much previous learner concordancing research (e.g., Bloch, 2007; H. Yoon, 2005) has maintained that learner variables such as inductive or deductive learning styles and tolerance to ambiguity can be important factors that influence the use and evaluation of reference resources. Investigating possible correlations between learning styles and reference resource consultation was not an aim of the present study, so I did not try to measure the learning styles of the participants directly using an inventory like the ones used in previous studies (e.g., H. Yoon, 2005). However, the participants’ survey and interview data along with observation of their i-Conc use brings into bold relief a pair of contrasting cognitive (learning) styles: intuitive (impulsive) vs. systematic (reflective) (Brown, 2007; Ewing, 1977). Ewing (1977) described an intuitive style as “making a number of gambles on the basis of hunches with possibly successive gambles before a solution is achieved” (p. 144) while defining a systematic style as reflecting on all the possible options before reaching a solution. The six participants can, to a certain extent, be placed somewhere on this intuitive-systematic continuum. Farthest on the intuitive side stands Yumee, who showed a strong tendency to perform a series of quick verification queries in her problem spaces, especially to confirm her “hunches” through SC queries without closely observing or reflecting on the results of each query. On the opposite side, Ian perused query results as much as possible to be sure of the appropriacy of the target item under consideration, so much so that he later often repressed his urge to do this and
purposely reduced the number of queries within a problem space to concentrate on the writing task at hand. As hinted at above, the difference in cognitive style was illustratively reflected in the proportion of SC queries to each participant’s total queries. To compare the two extreme cases, SC queries made up 46.5% of Yumee’s total queries while accounting for only 6.3% of Ian’s total.

Interestingly, these two different styles displayed by the participants also seemed to be related with a different set of **attitudes and levels of confidence** with respect to English academic writing. The participants closer to the intuitive end such as Yumee and Shia had in general relatively high confidence in their own writing, showed a greater tendency to trust their intuitions on L2 forms, and accepted query results more readily than did other participants. On the other hand, those closer to the systematic or reflective end such as Ian and Jae had relatively low confidence in their intuitions and tended to take cautious approaches by exploring alternatives suggested by the reference resources and critically evaluating their query results. While doing so, they tended to avoid the resources that they found less reliable in terms of data sources such as Google (see Figure 6.2 below). The differences in terms of attitudes and confidence were in turn clearly manifest in the proportions of satisfactory consultations, the actual success rate in each participant’s i-Conc use, and the types of reference resources consulted. To compare only Yumee and Ian again, Yumee’s satisfaction rate was 87.5%, the highest of all participants while Ian’s was 69.2%. However, the proportion of satisfactory consultations that resulted in an incorrect text formulation or revision was higher for Yumee (12.5%) than for Ian (3.8%). When it came to resources consulted, Yumee’s proportion of Google (unmediated, untagged and thus less reliable corpus) use (41.3%) was much higher than that of Ian’s (8.9%). Figures 6.2 and 6.3 compare the participants on these measures.

![Figure 6.2. Percentages of SC queries and Google use in each participant’s i-Conc use.](image)
Participants’ (meta-)linguistic knowledge also seems to have influenced their i-Conc use. Although the participants were all graduate students and advanced in their L2 proficiency, there were considerable gaps in the linguistic and meta-linguistic knowledge among the participants. While Ian and Goeun, the two language pedagogy majors, were the most knowledgeable, Jinho, the youngest and the only one who immigrated to Canada as an adolescent, was the weakest. It may be obvious that the writers with greater (meta-)linguistic knowledge would have greater control over lexical and grammatical items in their writing and thus probably encounter fewer problems in these areas of writing. Linguistic and meta-linguistic knowledge, however, may go beyond simply affecting the number of queries to affecting every step of problem solving processes. As documented in his case study, Jinho had difficulties throughout the different steps of i-Conc consultation related to his lack of (meta-)linguistic knowledge. First, not fully understanding the corpus linguistic concepts and the links between concordancing and lexico-grammar, provided in the tutoring sessions, Jinho used the concordancers like traditional dictionaries, carrying out many single-word queries for the purpose of verifying a match between a linguistic item and his intended meaning. Second, Jinho’s unfamiliarity with linguistic terms, especially names of word classes, made it challenging for him to represent a problem, devise a search strategy, and formulate a query in the ways required on resources like COCA. These difficulties he had, in turn, led to a high
proportion of unsatisfactory consultations (41.3%) and caused him to abandon i-Conc consultation in many of his problem spaces (21.7%). However, this lack of linguistic knowledge should not be equated with lack of language competence or writing proficiency, as can be seen in Jinho’s relatively high rate of successful problem resolution (78.2%). What he lacked was “explicit” linguistic knowledge such as familiarity with linguistic concepts and terms.

In addition to overall linguistic knowledge, exposure to and familiarity with academic/disciplinary register also appears to have affected the purposes for which the participants consulted the suite and their choices of resources. As explained in Chapter 3, GS and CSE were included to provide the participants access to domain- or discipline-specific usages and phraseologies, and in this study the queries for these purposes were coded as Reg (short for register). Overall, the four doctoral students (Jae, Yumee, Goeun, and Ian), though to varying degrees, performed Reg queries less frequently than the two MA students (Jinho and Shia) did. While Reg queries accounted for an average of 4.1% of all queries in the doctoral students’ data, it was 17.7% in the MA students’ data. The doctoral students also used GS and CSE less frequently than the MA students did. Some of them (Yumee and Goeun) explicitly mentioned that as they already were well aware of academic and discipline specific phraseologies through years of exposure to the topics in their individual fields of study, so they did not feel a particular need to attend to register in their written assignments. Moreover, even when they actually needed to verify the acceptability of certain word sequences they formulated in terms of register, GS and CSE often returned few or no matches. This pattern can in part explain why the under-use of these resources was more pronounced with the doctoral participants.

6.2 Text factors

Although the assignments the participants chose for the study were all academic papers, they varied in type, ranging from a research proposal to a conference paper and to a course paper. The participants’ data and their recollections reveal that the type and topic of writing also influenced their i-Conc use patterns. Course papers can vary in length and format as they are mostly defined by course instructors and contexts, and therefore can be less rigid in terms of structure, genre, and topic than other types of academic writing such as research papers or
proposals. Goeun’s self-chosen assignment was a course paper that set itself apart from other participants’ in that much of it was of a narrative nature, recounting her personal language learning experiences in prose that was often filled with emotions in a way that rarely featured in other participants’ written assignments. Goeun struggled to transform what had been mostly experienced and retained in her L1 (during her life years ago in Korea) into L2 text without losing subtle nuances and distinctions. She thus resorted to the bilingual dictionary for Equi purposes much more frequently than did other participants and carried out a number of SH queries (sentence/phrase hunting, see Table 3.8 for its description), which were only rarely performed by other participants. Other participants (Jae, Yumee, and Ian) also showed similar patterns in sections of their papers that contained L1 culture- and society-specific content carrying out Equi queries more frequently than in other sections. Figure 6.4 presents the proportion of Equi queries to each participant’s total queries.

Another potential influencing factor is **stages of the writing process**. The timing of the three screen recordings the participants had conducted while completing their writing assignments made it possible to examine differences in $i$-Conc use patterns between different stages of the writing process. Each participant conducted the first two recordings at the drafting stage and the final one at the revising/proof-reading stage. Although the boundaries were not always clear-cut for each participant, there were a number of major differences in $i$-Conc use patterns between these two stages commonly observed across all participants. First, they
performed a greater number of Equi queries at the drafting stage (16.1%) than at the revising stage (4.0%). Second, Bilingual (32.5%) was the most frequently consulted resource at the drafting stage while it was Google (27.2%) at the revising stage. These are not altogether surprising findings because the main writing activities differ at each stage—i.e., drafting a paper focuses more on transforming ideas into text while revising or proof-reading mostly involves confirming and polishing already-formulated text. These differences in focus were reflected most distinctly in the data of participants who followed relatively clear-cut stages of the writing process. Yumee, Goeun, and Shia showed considerably differing frequencies of problem and query types between the two stages of composing. The proportions of compensatory problems and elicitation queries were considerably higher at the drafting stage than at the revising stage. However, the rest of the participants, whose writing processes did not have such clear-cut boundaries between stages, did not show major differences in terms of problem and query types across the stages.

6.3 Context factors

The contexts in which the participants completed their self-chosen authentic writing assignments also varied widely on various levels ranging from the stage they were at in their studies to various demands on time from their personal lives. Whereas it is difficult to establish links between multi-layered context variables and specific i-Conc use patterns, there was one relatively clear context factor that emerged from the data as a major influence on i-Conc use patterns and the participants’ evaluation of the suite: each participant’s time availability for the writing assignment.

Jae, Yumee, and Ian had finished their coursework before participating in the study and their self-chosen writing assignments, a research proposal and conference papers, respectively, had no immediate fixed due dates. In contrast, Goeun, Jinho, and Shia took a number of courses while participating in the study, and the writing assignments they chose for the study had tight due dates that often overlapped with those for other writing assignments. Although all the participants struggled to juggle their busy personal lives and studies, the participants in the second group worked on their assignments under greater time pressure. This pressure, along with other factors discussed above, probably contributed to their relatively small numbers of
queries (an average of 72 queries as opposed to the first group’s 99) and shorter screen-recording times (see Table 3.7 in Chapter 3). On top of the physical number of i-Conc consultations done, the severe lack of time also adversely affected the participants’ i-Conc use in various ways in terms of consultation results and overall evaluation of the suite as writing assistance. For example, Goeun, despite her high language/writing proficiency, showed the highest rate of negative effects by i-Conc consultation of all participants (20%). These failed problem solving cases can be attributable not only to the factors described earlier, such as the overall greater difficulty of the type of problems she took up to solve using i-Conc, but also to the time pressure and fatigue she was suffering while doing the assignment. Some of the inaccuracies in her text were minor such as missing indefinite articles or plural markers, mistakes that she normally would not make. In turn, Shia was the one that explicitly linked “her absolute lack of time” to the quality of her i-Conc consultations. As described in her case study earlier, under severe time pressure, she showed a strong tendency to take up predominantly confirmatory problems (82.1%) and solved them through predominantly verification queries (79.2%), which required less cognitive exertion and therefore less time than the opposite types. Time pressure also took a toll on the rigor in her evaluation of query results. Although she was aware that it was not the right way to evaluate and apply query results, she accepted even a very small number of matches in the results as evidence of the acceptability of a given target item and then applied it to her writing. This behaviour also affected her evaluation of i-Conc as a writing supporting tool. While using the tool, she often felt it was a luxury she could not afford with the time pressure she was under.

6.4 Summary across cases

In summary, Figure 6.5 situates and compares the participants along these interacting factors. For particular factors, where a link could not be established with the participant, the cell is marked with “-”. In addition, for factors that exerted particularly strong influences on a given participant’s i-Conc use the cell is highlighted in shading.
| Factors                                                                 | Participants |
|---|---|---|---|---|---|---|
| Goals and expectations (performance vs. mastery)                      | P | P | P | M | P | M |
| Cognitive styles (systematic vs. intuitive)                           | S | I | I | S | I | S |
| Attitudes to (positive vs. negative) and confidence in their own writing (high vs. low) | -,- L | P, H | -,- L | N, L | -,- H | N, L |
| (Meta-)linguistic knowledge (high vs. mid vs. low)                    | M | M | L | H | M | H |
| Familiarity with academic/disciplinary register (high vs. low)         | H | H | L | H | L | H |
| Type and topic of writing (narrative, L2 culture)                     | -,- L | -,- L | -,- L | N, L2 | -,- L | -,- L2 |
| Stage of writing (little vs. much difference between stages)          | L | M | M | M | M | L |
| Time pressure (high vs. low)                                          | L | L | H | H | H | L |

Figure 6.5. Intervening factors influencing the participants’ i-Conc use and its evaluation.
CHAPTER SEVEN
DISCUSSION AND IMPLICATIONS

This chapter summarizes the major findings of the study in view of the research questions and then discusses what I make of them with reference to previous and future possible studies. I discuss a number of important issues about concordancing as a reference tool that have emerged from the findings. Specifically, the potential of online concordancing and other reference resources as a problem-solving tool is discussed in terms of the extent to which they can support a writer’s cognition and go beyond assisting with problem solving to facilitate language learning. This interpretation is then followed by implications and suggestions for theories, teachers, learners, learner training, and the further development of corpus-based reference tools and future research about them, arising from acknowledgement of limitations to the present research.

7.1 Major findings

The present study has been guided by the following five research questions: (a) how and for what do the participants consult i-Conc? (b) how does the use of i-Conc affect the participants’ problem-solving performance? (c) what are the strategies and challenges in the use of i-Conc? (d) how do the participants evaluate i-Conc as writing assistance, and (e) what are the possible intervening factors that influence individual differences (see Section 1.5.2 on page 10 for the full research questions). In the following six subsections, I present and interpret the major findings that address each of these five research question. The first research question is discussed in the first two subsections, 7.1.1 and 7.1.2.

7.1.1 Problems and queries

The participants used i-Conc, a suite of reference resources, both to test their intuitive hypothesis on L2 forms formulated based on their existing linguistic repertoires (i.e., for confirmatory purposes) and to extract linguistic items that were beyond their existing knowledge
or that were temporarily inaccessible due to memory lapse (i.e., for compensatory purposes). The actual queries they performed on i-Conc to solve these two types of problems were in turn prompted in two question forms: a *whether* question or *closed-form* question (verification), and a *what/how* question or *open-form* question (elicitation). As a group, the participants were shown to address more confirmatory problems and perform more verification queries during their writing assignments. This result confirms the findings of previous studies (H. Yoon, 2005; Park, 2010) in that the participants were shown to prefer using the corpora to confirm what they already knew rather than for discovering new facts about the target language. This overall trend can be in part attributed to the very fact that the participants used i-Conc as a reference tool rather than as a research (DDL) tool. In most DDL research, where the use of corpus tools were explicitly aimed to facilitate language learning through helping the learner to derive rules and regularities mostly by chance discovery, queries usually performed were predominantly of elicitation type. In contrast, in the present study, the reference resources were consulted to resolve immediate problems that arose while engaging in an authentic writing task, which in large part involved checking the acceptability of text segments already formulated in mind or on screen especially at the editing or proofreading stage of the writing process.

The participants as a group consulted i-Conc for a wide range of purposes, but more than 70% of their queries were carried out for four purposes: doing simple confirmation (SC, 24.3%), checking whether the target item delivers the intended meaning (Mn, 21.0%), finding or verifying a collocate (Coll, 18.4%), and finding an L2 equivalent (Equi, 10.9%). To look into only the primary (or initial) queries (from which motivations that prompted each i-Conc consultation can be inferred), the order changed as follows: Coll (21.2%), Equi (18.0%), SC (16.3%), and Mn (15.5%). This shows that SC and Mn queries were performed often as secondary queries to get corroborations for the primary query results.

To consider again only the top two purposes of the primary queries, collocation proved to be the most frequent purpose the participants consulted i-Conc for. Specifically, the participants searched for prepositional collocates the most frequently, followed by verbal and adverbial collocates. This result is similar to that of Park’s (2012) study, where finding and verifying appropriate prepositions and collocations were the most frequent purposes11. It can be

11 In this study, consultation purposes were classified somewhat differently with queries for lexical collocations and prepositions as separate categories.
interpreted that collocations, especially prepositional collocates, were the most problematic area for this specific group of participants. However, an alternative explanation would be that the concordancers were introduced to the participants as particularly effective tools for finding collocates during the training, and this step sensitized them to the problems of collocation and made them notice them more often than they normally would. This was indirectly verified by what some participants (Jae and Yumee) said in their interviews. One of the changes they experienced while using the suite was that they became more aware of and attentive to the lexico-grammatical aspects of vocabulary they used in their papers. This pattern echoes one of the main findings from H. Yoon’s (2005) study wherein participants raised their language awareness while using the corpus technology as a problem-solving tool.

Next, finding an L2 equivalent, or Equi, came out as the second most frequent purpose of the primary queries, suggesting that the participants, highly advanced L2 writers, working on an academic topic in their individual fields of study, to which they had been exposed mostly in English for years, still had to resort to their L1 frequently in conceptualizing content and to translate that into English. Analysis of the participants’ drafts and their stimulated recall protocols showed that Equi queries were in fact closely related to the type and topic of writing the participants did. Equi queries were shown to be carried out mostly in the sections of a narrative nature or where L2 culture-specific content was discussed. This result is one of the unique findings of this study in that most of previous concordancing studies could not capture these needs of L2 writers as they either did not provide along with corpus tools resources that could be consulted for Equi purposes or even when they were provided, the purposes they were used for were not examined. The only exception is Frankenberg-Garcia (2005), but the task done in that study was a translation task, which cannot be directly comparable to the types of academic writing the participants in the present study did.

In terms of linguistic categories, the participants’ queries were predominantly about lexical (45.6%) and lexico-grammatical (43.7%) concerns, which is consistent with Manchón’s (2011) analysis that the closer the writing task is to free writing, the more attention the L2 writer pays to lexis as opposed to grammar.
Turning now to the individual resources consulted, the participants as a group were in general well aware of what types of inquiry each resource was suitable for and so consulted them accordingly for distinctive purposes: Bilingual for Equi, COCA and JTW for Coll, Google for SC, and Monolingual for Mn, to list only the top purpose for which each of the major resources was consulted. To break down the participants’ queries by linguistic categories, the participants showed a tendency to perform queries of lexical concerns in the dictionary-type resources while consulting the concordancer-type resources more for lexico-grammatical and stylistic matters.

In terms of consultation frequency, Bilingual came out on top. This result was rather unexpected in that when designing the reference suite, this resource had initially been intended as just a supplementary resource, which had not been expected to be consulted much in academic writing while the training and feedback were heavily focused on the various, often novel, ways of solving linguistic problems using the concordancers. This tendency runs counter to the findings of previous studies conducted in similar settings (e.g., Park, 2010; H. Yoon, 2005). In those studies, dictionaries were available to use along with corpus tools while completing general academic writing tasks, and the participants developed over time a strong preference for the corpus tools they were provided with over bilingual dictionaries. In Frankenberg-Garcia (2005), the bilingual dictionary was the most preferred resource, but again, the task done in that study was a translation task, which by its nature required frequent consultations of the bilingual resource to find and confirm L1-L2 equivalents.

This frequent use of the bilingual dictionary in the present study, rarely observed or reported on in detail in previous research, can be attributed to a number of factors. First, as with most L2 learners (Tono, 2012), bilingual dictionaries were often the primary language reference resources that the participants had been using since they started to learn English and therefore they were most familiar with. Second, as noted in the previous section, the participants performed a great number of Equi queries, and Equi was the unique consultation purpose that could be served almost exclusively by Bilingual. There are a few other interesting factors that are closely related to the versatility of Naver, the specific online bilingual dictionary featured in the present study. Naver, developed by Korea’s biggest search engine provider with the same name, is not simply an online version of a paper bilingual dictionary, which is the case with most of freely available online monolingual learner’s dictionaries. But rather running on not
only multiple source dictionaries, but also a large database of Korean-English translation pairs of sentences, its functionalities go beyond the traditional bilingual dictionary to serve as a thesaurus and a parallel corpus. In addition, running on a powerful search engine, it is fast and stable but more importantly it allows multi-word searches and returns results in the concordance format along with the usual dictionary information of each constituent word. Taking full advantage of this versatility, the participants used Naver for the widest range of purposes of all resources.

Among the concordancers, COCA was consulted the most frequently and for the widest range of purposes, and most balanced in its use for both verification and elicitation queries, reflecting the variety of information that can be extracted from the resource with its query operators and options. However, this frequent use of COCA by the participants may not be an accurate indicator of their actual preferences. COCA was one of the new reference resources introduced to them, and they were encouraged to use as often as possible, so a novelty effect and also their awareness of being observed through screen recordings may have affected their choices of this resource. In fact, while all participants agreed that COCA was the most useful resource along with Bilingual, some criticized it for being too unwieldy with all its search options and complicated interface to be a handy reference resource to use while doing an actual writing task.

JTW, another concordancer in i-Conc working with a tagged corpus, proved to be a simple-to-use but powerful resource that provided a comprehensive range of collocational combinations in an organized way. Some participants (Ian and Goeun) found it especially learner-friendly and preferred it to COCA. However, after several consultations, other participants (Jae, Yumee, and Shia) simply stopped using the resource due to its unstable server and slow loading time. This trend suggests that the speed with which a tool can process a query and load the results can be an important factor that affects a writer’s choice of reference tools.

In this study, three Google search engines (Google, GS, and CSE), differing in size and scope of data sources, were included as concordancers. Well aware of the strengths (abundant data) and weaknesses (“dirty” and untagged data) of using a search engine as a language reference tool, the participants used Google predominantly for SC purposes, that is, confirming their hunches about L2 forms, and some participants even used it as their main resource. In contrast, GS and CSE were consulted the least frequently, accounting for less 5% of all resource
use respectively. This was another unexpected result, along with the frequent use of Bilingual, as these two search engines had been included in the suite specifically to meet the needs of graduate students’ disciplinary writing. During the tutoring sessions, the participants welcomed the inclusion of these resources in the suite for their potential for supporting their writing in terms of discipline-specific phraseologies, but in actual use they often became frustrated as GS and CSE, due to their small data size, returned no or few matches for queried items. This result does not necessarily mean that they are not useful resources for L2 writing. As some GALL (Google-assisted language learning) researchers suggest, these specialized Google engines can be great DDL research resources for investigating the differences in writing conventions between genres and disciplines such as referencing practices or types of reporting verbs (see Brezina, 2012; Hubbard, 2005 for examples). However, when used as a reference resource to find solutions to immediate problems that need to be solved to move ahead with the writing task at hand, the participants were not particularly motivated or able to conduct such investigations. As noted above, in this problem-solving context, the participants attempted to use these resources, instead, for quick confirmations on their hypotheses, and they proved to be ineffective as reference resources. The presence of Google Web next to them also contributed to their underuse. Google Web, one of the most familiar resources to the participants, was the only resource where, because of its sheer size, the participants could actually gauge the acceptability of their hypotheses against frequency information and readjust their queries accordingly. Realizing that GS and CSE did not allow much room for meaning negotiation because of their small sizes, most participants went straight to Google Web for SC queries.

Taken together, these findings provide some answers to the question “how corpus resources co-exist with online services like Google and online dictionaries,” asked by Perez-Padres, Sanchez-Tornel, and Alcaraz Calero (2012, p. 484). First, the present findings suggest that the use of concordancers as problem-solving tools while engaging in independent writing tasks may differ widely from when they are used as DDL tools for classroom tasks. The participants showed overall tendencies to (a) verify their existing knowledge more than to elicit new knowledge; (b) consult the bilingual dictionary more frequently than concordancers; and (c) prefer Google Web to more specialized search engines for quick confirmations. Second, the resources were shown to be mutually complementary. The participants overall consulted each concordancer for the purposes it was optimally suited for, and the participants strategically used
the different resources in combination within single problem spaces, especially in the form of getting a hint from one resource and expanding on it in another resource. These findings lend support to the arguments of some corpus-based language teaching researchers (e.g., Conroy, 2010; Flowerdew, 2009; Tribble, 2002) that corpus tools should be used in addition to, not in place of, more traditional resources. Third, however, as suggested by Naver, which has been evolving from a bilingual dictionary into a hybrid resource providing bilingual concordance lines, and also different types and functionalities of concordancers, the boundaries within and across resource types will get more and more blurred with further advances in computing power and text processing technologies.

7.1.3 Problem-solving performance

The participants’ overall effective use of i-Conc described above was also manifested in the relatively high rate of positive effects by i-Conc consultation on the participants’ actual problem solving results. In 69% of the problems they took up, i-Conc consultation led to a correct text formulation or revision. Although not directly comparable with the results from previous research because of the differences in task type, resources used and participants’ language/writing proficiency, this figure is higher than the corresponding results in Frankenberg-Garcia (2005) and in Park (2012), which were 65% and 54% respectively. In turn, i-Conc consultation resulted in an incorrect text in 13% of the problems that participants addressed, while in 18% of the problems, the participants abandoned i-Conc consultation midway and solved the problem on their own. By problem type, compensatory problems proved to be more difficult to solve by using the reference suite than confirmatory problems, which showed a lower rate of positive effects by i-Conc use but a higher rate of i-Conc consultation abandonment. This trend reflects possible differences in the degree of cognitive exertion and (meta-)linguistic knowledge required in consulting i-Conc to solve these two problem types. Confirmatory problems were simpler in terms of both query term formulation and evaluation of query results. The resolution of a typical confirmatory problem involved performing a verification query, that is, simply querying a word or word sequence in question in a resource and checking in the results the presence or frequency of its instances or its match with the intended meaning. Compensatory problems, by contrast, were typically solved through
elicitation queries, which usually required the participants to strategically use query syntax/operators or choose a specific option and to pick out the most appropriate choice from multiple alternatives provided in the results. By the content of the problems, i-Conc consultation was shown to be the most successful for checking simple grammatical points such as transitivity of a verb (argument pattern, AP) and noun endings (NE). On the other hand, the rate of positive effects by i-Conc was particularly low for the problems of finding alternatives; in a great proportion of these problems (both uAlt and sAlt), the participants abandoned their i-Conc consultation midway (about 44% and 53% respectively).

Meanwhile, there were broadly two major causes for the negative effects by i-Conc consultation. One was misinterpretation of the query results where the participant took a wrong item or false confirmation from the results. The other was misapplication of the solution found to writing, involving cases where even though the participants found the right solution through i-Conc consultation they produced a grammatically incorrect or stylistically inappropriate text. This pattern can again be attributed not only to the participants’ overall language proficiency but also to the pitfalls of i-Conc consultation into which the participants were observed sometimes to fall, as is summarized in the next section.

As with the patterns of i-Conc use in terms of problems addressed, queries performed, and resources consulted as demonstrated in Chapters 5 and 6, there was a wide variability in the problem solving performance, which cannot be explained by one single factor. For example, Goeun and Ian, the two most advanced writers, showed a below-average performance respectively in terms of positive effects by i-Conc use, and also the highest rate of consultation abandonment. While this trend confirms that higher proficiency does not necessarily lead to better performance with a corpus tool (Boulton, 2010b; Yoon & Hirvela, 2004), it can be in part explained by the severe time pressure they worked under and more importantly the fact that as more advanced L2 writers they tended to take more risks trying to deliver subtle nuances using more complex lexical phrases and syntactic structures (Paquot & Granger, 2012), which led to frequent errors (Goeun) or abandonment (Ian).

7.1.4 Strategies and pitfalls
As noted above, overall, the participants used the reference suite quite effectively with i-Conc use contributing to correct text formulation/revision in about 70% of their problems. To a large extent, this effective use was made possible through strategies the participants employed or newly developed as they became more familiar with the suite. These strategies can be broadly broken down into query and evaluation strategies.

Query strategies are ones that the participants used at the early stages of the i-Conc consultation sequence (see Figure 4.2 in Chapter 4) when formulating and refining a query. First, when formulating queries, some participants often used highly efficient strategies that involved (a) performing a preliminary query in one resource and expanding on it in another resource, (b) starting off a confirmatory problem space with an elicitation query, rather than the usual verification query, or (c) using search operators creatively. Second, when initial query results were not satisfactory, the participants often performed further queries by strategically refining the initial query term. They replaced (part of) the initial query term with an alternative in terms of meaning, POS, or syntactic structure. Or they used the strategy of making the initial query more specific by inserting a concrete word, or a grammatical marker, or more general by reducing the number of words or using more general words.

Evaluation strategies, in turn, refer to the strategies the participants used to evaluate query results and apply the solution found to their writing. The most frequently used strategy of this type was checking frequency information, and then choosing the most frequent item from multiple alternatives (in an elicitation query), or accepting a certain number of tokens/hits of the target item as evidence of its acceptability (in a verification query). Some participants used, though much less frequently, their familiarity with the items or authority of their data sources as criteria for choosing the item to apply to their writing. Finally, while performing a query, the participants often picked out the words and phrases that they found by chance that were not directly related to the item being queried, and they then used them in their writing.

In comparison with the participants in previous studies that reported on strategy use (e.g., Kennedy & Miceli, 2010; Park, 2010) in reference resource consultation, most participants in the present study were shown to be effective on the whole in their use of resources in terms of the number of strategies and frequency of their uses. This trend may also be attributed to multiple factors. First, among others, unlike the resources provided in most previous studies, i-Conc features multiple resources that differ considerably in their functionalities and querying
methods. This choice of options posed some challenges but it also encouraged the participants to employ appropriate strategies for different resources and to further develop strategies to effectively use multiple resources within a problem space. Second, in addition to the participants’ overall high levels of language proficiency and cognitive skills, the training they received seemed to help. Unlike in previous studies, most of which were conducted in classroom settings, each participant in the present study was provided with one-on-one personalized tutoring in basic strategies at the beginning and also received contingent feedback on ways of improving their i-Conc use at the end of stimulated recalls.

Despite their overall effective use of i-Conc, the participants were also observed to fall into pitfalls in its use, which contributed to incorrect text formulation or revisions, false perceptions of successful problem resolution, and/or consultation abandonment. There were three major pitfalls identified. First, lack of rigor in observing query results was the most common cause for unsuccessful problem solving outcomes. The participants sometimes simply picked the most frequent item or checked the target item’s frequency information but did not check whether it carried the intended meaning or was appropriate in the intended context. In other cases, they only focused on the form of the target item and did not notice the specific syntactic environment in which it occurred and so ended up producing an ungrammatical text. The second pitfall commonly observed was lack of flexibility in trying out different options. The participants missed out on opportunities to reach better solutions or the same solutions more effectively by using only certain ways of representing problems and using only certain resources and query options. The last pitfall stemmed from the limitations of the resources themselves. Each resource in i-Conc had a set of limitations as a reference resource, which often got in the way of the participants’ problem solving. For example, Google sometimes returned a considerable number of hits for even an ungrammatically formulated query term, giving the participant false confirmation on that queried word (sequence).

These pitfalls were to some extent the flip side of the strategies the participants employed for successful i-Conc consultation. It was sometimes inevitable for the participants to fall into these pitfalls when using (and mis- and over-using) the strategies. For example, using frequency information was one major strategy the participants used in evaluating query results. Under severe time pressure, however, the participants often used the strategy by default but could not afford to check the match between the target items and the intended meanings.
Another example can be found in the routinized consultation patterns developed by the participants to increase their efficiency. One such pattern was running a series of quick SC queries on Google. Some participants (Yumee and Shia) towards the end of the study displayed great speed and efficiency in using this consultation pattern. However at the same time they were so fixated on this specific set of query type and resource that they largely left untapped much of the potential of the suite as a reference tool.

7.1.5 Participants’ evaluations of i-Conc

The participants’ evaluations of i-Conc as writing assistance was positive overall, and they said they were willing to use i-Conc, or at least some of the resources in it, for their future English writing. As to the advantages of using i-Conc for English academic writing, the participants’ responses were along the lines that i-Conc provided them with easy access to mutually complementary reference resources in a single interface, which made it possible to get corroborations on the linguistic items in question from different resources relatively quickly and therefore gave them reassurances that they could use them with a certain degree of confidence in their writing. There were other important benefits of using i-Conc mentioned by, if not all, the participants. To list only the major ones, first, the presence of multiple resources that can easily be accessed motived the participants to venture beyond their current linguistic repertoires and experiment with new ways of encoding intended meanings. Second, the participants became more attentive to the accuracy and approproacy of form-meaning mappings they did while writing and increased their awareness of the lexico-grammatical aspects of language they used. Third, all these factors combined to increase, if not drastically, their confidence in the language aspects of English academic writing, providing them with a sense of independence as L2 writers.

However, the participants also experienced various difficulties and challenges in using the reference suite, especially the concordancer-type resources. As in most previous studies, the first and foremost challenge that all participants shared was the time-consuming nature of i-Conc consultation. Most participants expressed frustration with the time taken to go through the consultation cycle from devising a search strategy to formulating a query term to evaluating the results to find what they were looking for. Some participants found going through the consultation cycle—especially when using concordancer-type resources—not only time-
consuming but also cognitively taxing as it required them to make judgments and decisions at every step, when much of their attentional resources were already consumed by other various aspects of the writing task at hand. Closely related to this additional cognitive load, some participants also reported that they sometimes became so engrossed in i-Conc consultation that they got distracted from the writing task at hand and had their flow of thoughts interrupted. For some participants, i-Conc consultations, if not always, created a cognitive overload and distraction for them rather than the extension of knowledge resources it had originally been assumed to be.

A closely related issue that some participants considered a shortcoming of i-Conc was the nature of interactivity they had with the suite. The participants pointed out that unlike getting support from a human tutor/expert, it is still almost solely the user’s responsibility to find a solution by feeding the right query terms into the resources and making the correct judgments about the results on her own, and even then there is no guarantee that the solution found is the right one. In this sense, i-Conc consultation was to these participants still more of a unilateral rather than authoritative, dialogic activity with no division of “responsibility” for the text being constructed/revised, which they usually had when working with human teachers and tutors.

These negative reactions suggest that the participants, if not all, perceived i-Conc, especially its concordancer-type resources, differently from what they had been intended to be. While the results overall lent support to the theoretical assumptions that concordancing can be a cognitive tool that amplifies the user’s cognitive power in a cognitive division of labor, as discussed in Chapter 2, some participants did not perceive it as such but instead considered it cognitively disruptive. In addition, while welcoming the notion of greater autonomy at the beginning of the study, some participants were later shown to struggle with the uncertainty of the solutions they found on their own through i-Conc consultation. Specifically, some of the negative reactions by the participants seem perplexingly at odds with the conclusions and findings from previous studies. For example, the greater responsibility that the participants felt that came with more judgments and decisions to make using i-Conc was in fact discussed as one of major benefits of using corpus tools in H. Yoon (2008). Interactivity, which these participants thought i-Conc lacked, was what Park (2010) argued was the major strength of the corpus tool used in his study. There appear to be large, possible gaps between teachers’ or educators’ perspectives and learners’ perspectives in these respects.
However, to take a step back and put the participants’ evaluation of *i*-Conc in perspective, these negative reactions can be seen as the flip sides of the very things that some participants considered benefits of using *i*-Conc. While some participants saw *i*-Conc as enabling them to go beyond the usual vocabulary and syntactic structures they normally use, and to be more autonomous in their L2 writing on the one hand, but other (and even the same) participants felt that greater attention to the language aspects of their writing was distracting, and they were not sure how to deal with the uncertainty regarding the solutions they found in their own consulting *i*-Conc, on the other hand. In previous studies, only the benefits of using corpus tools—e.g., greater responsibility for one’s own writing, increased autonomy (H. Yoon, 2005, 2008), and the cognitive division of labor afforded by dialogic interactivity with the tool (Park, 2010)—were emphasized, but little was reported on participants’ mixed feelings and even possible resistance to these notions.

By contrast, that other side of the coin was brought into sharp relief in this study, perhaps because of the broad educational context in which the participants were situated and the unique research settings of the present study. Broadly, the participants were all graduate students whose approaches to academic writing for course assignments typically focused on meaning or content. The presence and use of *i*-Conc in this study heightened the participants’ awareness of language problems and provided a means to address them as they arose. Sometimes these affordances offered by *i*-Conc led the participants to focus on form (i.e., linguistic features of writing) more than they normally would. It was on these occasions when the main goals of the participants and the affordances of the tool were misaligned and thus *i*-Conc consultation felt like an extra cognitive burden. This trend was uniquely observed in the present study in contrast to most of the previous studies in which learners’ concordancing activity was conducted as a main task as part of a classroom curriculum for language or writing courses. Secondly, all participants but one had used writing centre services extensively, and so they were used to writing centre consultants “unilaterally” giving them “authoritative” solutions to their problems, which in turn gave them a degree of certainty about their language use and writing skills.

At the more micro level, the specific research settings of the present study added another layer of complexity to deal with for the participants. Firstly, unlike in most previous studies where only one or two reference resources were provided, the present participants had to consult
the multitude of hitherto-unfamiliar resources in i-Conc. With the exception of one participant, they had never used many of the resources in the ways encouraged by the researcher. Although initial training and informal feedback was provided at the end of each stimulated recall, most of the time the participants were left to their own devices without ongoing guidance about how to turn the extra cognitive load to their advantage. Secondly, this feeling may have been exacerbated by their knowing that their uses of resources were closely observed in short spans of one-hour screen-recordings, which caused some participants to feel pressured to consult the suite even when they did not feel a particular need to.

This interpretation suggests that L2 writers’ perceptions and evaluations of i-Conc and similar computer-mediated tools may vary depending on the context for which they engage in writing and their needs and goals arising from that context. Manchón and Roca de Larios (Manchón, 2011; Manchón & Roca de Larios, 2007) classified the educational purposes of L2 writing research into three broad strands: *learning-to-write*, *writing-to-learn-content*, and *writing-to-learn-language*. Applying these three purposes to L2 writing contexts, the reactions documented from these graduate-student participants while producing assignments for their degree programs can be related to purposes of *writing-to-learn-content*. A similar study in a *writing-to-learn-language* context—such as an EFL course or an ESL academic writing classroom where writers are motivated and guided to attend to the linguistic features of their writing—may well produce different and perhaps more positive reactions to i-Conc.

### 7.1.6 Individual differences and intervening factors

The six case studies showed wide variability in almost every aspect of the individual participants’ i-Conc use patterns and evaluations thereof. Cross-case analyses identified salient factors that may have contributed to this variability across the participants. The majority of the identified factors were writer-specific ones that included the participants’ overall goals and expectations for their own academic writing, cognitive styles, attitude to and confidence in L2 writing, (meta-)linguistic knowledge, and familiarity with disciplinary registers. There were also text- and context-specific factors. The type of the writing assignment and the stages of the writing process were the two major text-related factors while time availability emerged as one clear common contextual factor influencing the participants’ referencing behavior and patterns.
Individual participants’ specific i-Conc use patterns and evaluations of i-Conc as a writing supporting tool appear to be linked to the interactions of these intervening factors, each of which had varying degrees of influence depending on the participant and situational context. For example, a participant’s frequent use of verification queries, one of the common i-Conc consultation patterns, should be seen as a result of any combinations of the following, but not limited to, reasons rather than linked to a single factor. It could be that the participant’s cognitive style was more intuitive than reflective and therefore she tended to test her intuitive hypotheses more often than to seek alternatives suggested by the suite. It may also be that she perceived i-Conc as a proofreading tool and so used it mostly to seek confirmations on already formulated texts. Or she simply could not afford to perform more time consuming, cognitively laborious, types of queries under severe time pressure. In other words, the participants’ individual traits, goals, and ongoing needs arising from the macro- and micro-contexts in which they were writing interacted to affect how they perceived and acted on particular affordances of i-Conc.

7.2 Implications

7.2.1 Implications for theories: i-Conc as a cognitive tool

The main theoretical assumption for the present study was that linguistic reference resources, especially hands-on concordancing, can provide L2 writers with an intellectual partnership and extend their knowledge and thinking while problem solving. The results of the present study showed that overall i-Con can be an effective cognitive tool that enhances the L2 writer’s on-the-fly problem solving. However, there was a wide difference in the effectiveness and efficiency with which individual participants made use of the tool. The role of i-Conc as a cognitive tool did not always lead to successful problem-solving performance nor was it always received as extensions of the participants’ cognition. In what follows, on the basis of the results presented above, the cognitive effects of i-Conc use are explored in greater depth.
7.2.1.1 Cognitive effects of i-Conc

Returning to issues discussed in Chapter 2, Jonassen and Reeves (1996) defined cognitive tools as “technologies, tangible or intangible, that enhance the cognitive powers of human beings during thinking, problem solving, and learning” (p. 693). Seen from the observations and the participants’ own recollections of their i-Conc use, the reference suite seems to fit this definition nicely. How i-Conc actually worked as a cognitive tool can be discussed on two dimensions. Educational psychologists studying cognitive tools suggest that the effects of these tools can be broken down into two kinds (Jonassen & Reeves, 1996; Salomon et al., 1991). Effects with a technology refer to those effects that the tool has when the user works in an intellectual partnership with it and enhances her intellectual performance. Effects of a technology, in contrast, refer to “cognitive residue” (Salomon et al., 1991, p. 2) that stays with the user after using the tool.

To look at the results for these two kinds of cognitive effects, first, i-Conc was shown to serve as an “intellectual partner” to the participants through cognitive division of labor, and in large part enhanced their intellectual performance for solving immediate problems that arose while writing. To give only the simplest description of the typical division of labor that happened in the participants’ problem solving, i-Conc provided the participants with lexical and grammatical options that they would not have been able to access instantly as items were not either acquired yet or fully internalized (i.e., compensatory problem solving); i-Conc also made it possible to test and readjust their intuitive hypotheses about form-meaning mappings (i.e., confirmatory problem solving). While the tool retrieved, sorted and presented the needed information, the participants, for their part, focused more on higher-order cognitive processing such as making a judgment based on the provided information about the accuracy and/or appropriacy of the linguistic item at hand and then making a choice for a further action. Thus, effects with using i-Conc occurred in the form of distributed cognition while jointly solving linguistic problems in ways that the participants’ unaided minds alone would not have been able to do, at least not as effectively.

Effects of using i-Conc, in turn, cannot be observed directly from the screen recordings or the participants’ recollections of their i-Conc use as they are, by definition, cognitive effects that the user is left with after using the tool. However, they can be, to some degree, inferred from the participants’ evaluations of the tool, especially the changes they said they experienced.
while using the reference suite. These include, but are not limited to, greater sensitization to linguistic problems while composing (e.g., greater attention to the accuracy and appropriacy of form-meaning-function mappings and increased awareness of lexico-grammatical aspects of phraseologies), improvements in search skills and strategies, and most importantly, potential language acquisition. These changes reported by the participants may constitute the “cognitive residue” that i-Conc use left that can be transferrable to other problem solving situations in which the reference suite is not accessible or different types of tools are used.

However, as noted a few times already in the previous sections, most of the participants also experienced varying degrees of frustration with respect to the cognitive processes that were involved in using i-Conc. While the reference suite was supposed to amplify their cognition in problem solving, ironically they sometimes found the use of the tool cognitively draining and even distracting. This irony is explored in depth in the next section.

7.2.1.2 Cognitive extension vs. cognitive overload/distraction

The possible causes for the participants’ occasional struggle for and resistance to the extra cognitive load involved in using i-Conc have been discussed in the previous sections with particular reference to contextual factors. In addition to them, I suggest here more general explanations drawn from the very nature of cognitive tools and the writing process. First of all, a cognitive tool is a knowledge construction or facilitation tool that is by definition supposed to increase mental activity and facilitate meaning making processes on the part of the user (O’Sullivan, 2007). Even without i-Conc, the participants’ cognition would have actively worked to solve lexical and grammatical problems they noticed while writing their papers. However, the use of i-Conc itself required a new type of cognitive effort that entailed search strategy formulation, search results observation, and decision making. This complexity may have collided with the overall expectations the participants had for the suite when it had first been introduced to them. As hinted at in the final interviews with Jae, Goeun, and Ian, the participants might have expected a tool with artificial intelligence that would work for them and take most of cognitive load off them when finding elusive words and structures for them, but not a tool that they should work with that therefore required their active participation in the partnership in terms of planning and decision making. This largely unexpected extra cognitive load became a source of frustration for the participants.
Secondly, cognitive processing while writing involves allocating attentional resources to different aspects of writing such as content, organization, and language (Bereiter & Scardamalia, 1987; Cumming, 1990; Flower & Hayes, 1981; Torrance & Galbraith, 2006). How much of the available attentional resources the writer allocates to each aspect may vary widely depending on the language and writing proficiency of the writer, the writing task, and the stage of the writing process. Many of the participants found i-Conc consultation particularly disruptive at the drafting stage, wherein they tended to focus predominantly on content (what to say) rather than on rhetorical concerns (how to say it). As noted earlier, however, the use of i-Conc, encouraged the participants to pay more attention to the language aspects of their texts than they normally would while drafting a paper. While doing so, they sometimes became so engrossed in finding solutions to local problems that they got distracted from the global content of the text and had their flow of thoughts interrupted. This was when they felt the use of i-Conc distracted rather than extended their cognition.

However, one may ask whether the increase in mental activity and greater attention to the language aspects of text brought on by the use of the reference suite should be seen as negative effects. They are in fact what would be encouraged in a typical DDL classroom, where language learning is an explicit goal. Although language learning was not one of their explicit goals for doing the self-chosen writing assignments, the increased exposure to language data and greater attention to the lexical and grammatical aspects of their texts may have incidentally facilitated participants’ language acquisition in the process of problem solving, as some participants themselves suggested in their final interviews. Viewed in this way, the cognitive overload or distractions that often frustrated the participants may be an important benefit in disguise, of which they were not aware. Distinctions are often made in theories of second language acquisition (SLA) between learning/acquisition and uptake and between retention and uptake (Manchón, 2011), suggesting that the successful resolution of a linguistic problem should not be equated with the acquisition of that linguistic item because the user may formulate a correct text applying the solution she finds without fully internalizing it. However, the ways in which the present participants interacted with i-Conc can be seen to amplify the potential of L2 writing for facilitating language acquisition.
7.2.1.3 Potential role of i-Conc for language acquisition through L2 writing

In a DDL language classroom, where deriving rules and regularities of the target language is the main goal and task, concordancing is directly used as a tool for language acquisition. However, when concordancing is used as a problem solving tool for L2 writing, which was the case with this thesis study, reference resources may play a secondary role that maximizes or amplifies the potential of L2 writing for facilitating language acquisition.

A number of SLA theories lend support to the facilitative effects of L2 written production for language learning (see the special issue, Volume 21(4), of the Journal of Second Language Writing: Exploring L2 Writing-SLA Interfaces, for a comprehensive review). To begin with, the very nature of writing itself provides a facilitative condition for language acquisition. In comparison with spoken language production, the permanence of written text and the relatively slow pace of language production provide more opportunities than, for example, during conversation to attend to language and thus greater expectations for its accuracy and appropriacy (Chapelle, 2008; Cumming, 2011; Williams, 2012).

Next, many think-aloud protocols conducted in previous L2 writing research (e.g., Cumming 1990; Roca de Larios et al., 1996, 2006) suggest that the cognitive processes the L2 writer engage while writing may be similar to the process of meaning negotiation in which the language learner dialogically adjusts her production based on an interlocutor’s feedback, which has been suggested to drive language development (Long, 1983). “Self-clarification and linguistic adjustments” (Cumming, 1990, p. 503) that can be observed from the decision making processes of L2 writers resemble feedback and repairs during verbal interactions. The differences from oral interactions are that while writing, meanings are negotiated intrapersonally or mentally (Cumming, 1990; R. Ellis, 1999).

The Output Hypothesis (Swain & Lapkin, 1995; Swain, 1998), in turn, argues that while producing the target language, the learner can notice gaps in their linguistic knowledge and that this noticing of linguistic problems and subsequently addressing them triggers cognitive processes involved in second language learning. Such noticing is triggered by feedback the learner receives during and after language production in the forms of external feedback from an interlocutor or reader and internal feedback, which resembles the intrapersonal meaning negotiation described above (Hanaoka & Izumi, 2012).
In SLA theories, interaction and feedback play a crucial role in setting a stage for language acquisition and development. This may be where a language learning role for i-Conc becomes clear. As can be seen in the participants’ multi-query interactions with i-Conc (see Figure 4.3 in Chapter 4 for an example), the suite allowed the participants to engage in meaning negotiations rigorously, serving as a more competent interlocutor who provided external feedback extracted from attested language data. As some participants suggested, i-Conc is far from a more competent human interlocutor (i.e., tutor or teacher) who can dynamically assess the learner’s level and then adjust their level of scaffolding. However, insofar as i-Conc served a mediating tool with which the participants could seek solutions to their self-initiated problems by testing their own hypotheses and accessing the linguistic items that are beyond their linguistic repertoires, it may reduce the extent of guesswork involved in intrapersonal meaning negotiations and compensate for the lack of external feedback in L2 learners’ unaided writing. In addition, feedback from reference resources is provided immediately through in situ negotiations between the user and the reference tools so this immediacy of feedback may lead to successful learning (Frankenberg-Garcia, 1999).

Finally, the Involvement Load Hypothesis (Hulstijn & Laufer, 2001) suggests that the use of i-Conc for L2 writing may also directly enhance language acquisition. In vocabulary acquisition, the hypothesis claims, tasks with high values for need (the need to know or use a word), search (a search for the right word to express the intended meaning), and evaluation (the comparison of different alternatives to choose the most suitable selection) promote vocabulary acquisition more than those with low values for these features. These three components correspond to the typical problem solving sequence identified in the present study, which consisted of noticing a problem (need), querying (search), and evaluating the query results (evaluation). These qualities suggest that i-Conc consultation sequences may enhance the likelihood of acquisition of lexical items. Specifically, the use of i-Conc while writing may be linked to two modes of language acquisition (Nobuyoshi & R. Ellis, 1993): It may have helped the participants to internalize new forms (through compensatory problem solving) and to have a better control of already internalized forms (through confirmatory problem solving).

As presented in Chapter 2, the language learning potential of concordancing has typically been discussed from the perspective of DDL and so concordancing research has also been largely confined to DDL settings. However, the links made above between SLA theories
and the ways the participants interacted with i-Conc imply that the use of concordancing (and other similar resources) as a reference tool may go beyond simply short-term problem-solving, to potentially facilitating long-term language development as suggested by the SLA theories reviewed above.

7.2.2 Pedagogical implications

The findings of the present study have a number of important pedagogical implications for learners, teachers, learner training, and the further development of corpus-based reference tools.

7.2.2.1 Implications for learners

Even advanced learners still need support in the language aspects of their writing because the external support that they can get is not sufficient and seldom is immediate. As one of the present participants poignantly observed (see Jae’s case study in Section 5.1), this lack of support, in both amount and immediacy, for lexical, grammatical and other surface-level problems can be a major source of frustration for L2 writers at North American universities. In this sense, the addition of concordancing tools to L2 learners’ usual reference resources can enhance their ability to self-edit and proofread at least the surface levels of their texts by providing them mediations with which they can test their intuitive hypotheses rigorously and venture beyond their usual linguistic repertoire with confidence.

In addition, as discussed in detail in previous sections, the greater attention to language features of writing, and active meaning negotiations that come with the use of a reference suite during writing are likely to promote L2 writers’ language learning, and genre-specific writing proficiency, particularly in the form of heightened lexico-grammatical awareness. This benefit, although significant for all levels of L2 learners, may be especially useful and needed for advanced L2 writers like the participants in the present study. Advanced L2 learners, especially graduate students who take mostly content courses and get little support for the language aspects of their writing, tend to have little opportunity to engage in explicit language learning and reflect on their own language use. While these students may not particularly link their writing assignments with language learning, multi-resource reference tool consultation while
completing those assignments would enable them to notice and address their own written errors and fossilization in their language use and to try out new (or more appropriate) ways of expressing their ideas suited to the registers and conventions of their academic disciplines.

7.2.2.2 Implications for teachers

The findings of the study also have important implications for L2 writing pedagogy and writing teachers. Firstly, students’ autonomy enhanced by the use of a reference tool can also be a major benefit for writing teachers and language support staff in universities, who often struggle with a lack of resources and time or a lack of flexibility under curricular or policy restrictions (e.g., writing centres refusing to surface-edit or proofread student papers) especially in dealing with lexical and grammatical errors in their students’ texts. By encouraging L2 writers to solve language problems on their own by making effective use of reference resources, these writing teachers or consultants may relegate some of their responsibilities to students themselves and be able to focus more on other aspects of students’ writing such as the organization of information, ideas, and discourse-level concerns (Conroy, 2010; Gaskell & Cobb, 2004; Milton, 2006; Park, 2010). This implication does not mean that writing teachers can be relieved from their responsibility of attending to students’ lexical and grammatical problems, but as many corpus-based language learning scholars have argued (e.g., Bernardini, 2004; Chambers, 2007; Gavioli, 2009), teachers necessarily should play a role of facilitators who monitor and support students’ problem solving with reference resources.

To play that facilitator role, however, writing teachers themselves must be knowledgeable in concordancing and other useful reference resources. The reality is that corpus tool use (and the combined use of different references) has not been popularized in language classrooms but is still practiced largely on an experimental or exploratory basis, so most writing teachers and consultants themselves may not be familiar with new types of reference resources. For this reason training should be provided for teachers as part of pre-service and in-service teacher education programs (Breyer, 2006; Chambers, 2007; Flowerdew, 2010; O’Keefe & Fiona, 2003).

A related point is that the mixed methods design of the present study, especially the use of the participants’ screen recordings of their writing processes and stimulated recalls of these processes, proved to be effective in showing what types of problems L2 writers notice and
address (or not) with reference tools during writing, at least at the lexical and grammatical levels, and to what extent what they intend to do matches what they actually do in terms of lexical searches and syntactic encoding. This type of data can provide valuable input to inform the development of scaffolding appropriate for individual L2 writers and the overall foci of writing instruction for specific groups of learners. For example, for the six participants in the present study, it would be helpful if a teacher were to provide detailed instructions on the concepts of lexico-grammar (inseparability of lexis and grammar) including collocation and other formulaic language and also on effective uses of students’ L1 while performing L2 writing.

7.2.2.3 Implications for learner training

What specific forms should pedagogical applications take to realize these potential benefits while minimizing their drawbacks? First of all, the findings of previous studies (e.g., Kennedy & Miceli, 2001, 2010; H. Yoon, 2005, 2008) and of this present study confirm that it is a lengthy process for L2 writers to master using reference resources as problem solving tools independently. Gradual and guided training and meaningful engagement with reference tools are required over an extended period of time. Independent use of concordancing and other reference resources can be better practiced if preceded by a general understanding of basic concepts around concordancing and reference-resource consultation in general, and sufficient hands-on practice with ongoing guidance from teachers or tutors. In this respect, concordancing and its combined use with other resources could usefully first be introduced into academic writing courses for students at intermediate levels of English proficiency. As noted earlier, compared to advanced L2 writers like the participants in this study who were oriented in tasks for their graduate studies to writing-to-learn content, (Manchón, 2011) or writing to mean (Byrnes, 2013), intermediate-level L2 learners, while having levels of (meta-)linguistic knowledge needed for concordancing, are more likely to engage in writing to learn the target language. For this reason, using reference resources for lexical and grammatical problem solving should help improve such students’ writing directly while causing less cognitive disruption than was seen with the present participants. One important insight from this study is that learners should be guided to build a meaningful cognitive partnership with reference tools on their own by learning over time to come to terms with and eventually turn to their advantage the cognitive exertion
involved in using reference tools as they move across different writing contexts (e.g., *writing-to-learn content*, or *writing to mean*).

At the same time, as the present case studies clearly demonstrated, even within similar educational contexts, individual L2 writers’ processes and outcomes of reference resource consultation can vary widely. Thus, special care and guidance should be given to individual needs of students through one-on-one conferences that may resemble the stimulated recall and feedback sessions conducted in this study. Then, over time when students are increasingly familiarized with reference tools, and the different types of problems that they can tackle with the tools, students can be progressively guided to use these tools outside of the classroom for their own content-course writing assignments.

For those L2 writers who do not take language or writing courses at universities, support programs such as writing centres can offer classes and workshops that introduce concordancing tools and other useful reference resources for L2 writers. Further, writing centre tutors may solve together with students the language problems commonly found in their papers using concordancing and other useful resources in their one-on-one sessions, through which the students learn the functions of different resources and strategies for using them effectively.

Then, what should be the content of this learner training? Advances in computer technology and reference resource research have and will come with new types of reference resources and changes in the scope and functions of existing types of resources, so it would be important to start learner training with general principles of effective reference resource consultation for problem solving that can apply to different types of resources universally. These general principles can in turn include specific attitudes or perspectives that learners need to cultivate to have meaningful engagement and a productive cognitive partnership with reference resources on the one hand and general strategies for effective searches and evaluation/application of search results on the other.

To discuss facilitative attitudes for successful reference resource consultation for L2 writing first, I present three perspectives distilled from the findings of the present and previous studies (e.g., Kennedy & Miceli’s 2010 “apprenticeship” approach). The first perspective is the recognition that it can be a long process to become familiarized with and reach a certain level of mastery of a reference tool. Only through actual use, practice, and some trial and error can users discover functions that meet their specific needs while writing in their target language and so
develop their own strategies for using each function. Some participants in the present study were observed to give up using certain resources due to their initial steep learning curves before developing a feel for their unique functions and how they could support their writing. Therefore, L2 writers should invest a certain amount of time and effort in figuring the resource out on their own and “making it [their] own” (Kennedy & Miceli, 2010, p. 38). One should be reminded that overall effective use of online/offline dictionaries and search engines such as Google are the results of years of experience and repeated use.

Second, L2 writers should be aware of the potential of reference resource consultation for their own language acquisition. When using a reference tool while engaging in a writing task, the writer’s main intention may be to use it as a problem-solving tool to complete the task at hand as efficiently as possible with no particular intention of language learning. However, reference resources, when strategically used, can help writers to do rigorous meaning negotiations that are based on attested language data and provide immediate external feedback, therefore allowing writers to access the linguistic items that lie beyond their usual linguistic repertoires. In other words, L2 writers should realize that reference resource consultation done for problem solving lead to their language development in the sense of learning more about the language and using the language better. In this sense, the extra cognitive load involved in the use of reference resource consultation should be accepted as part of the learning process, as long as it is not too heavy to handle under the given constraints. Even the cognitively disruptive nature of reference resource consultation during drafting can be turned to advantage. As some of the present participants reported, greater attention to and uptake of linguistic problems at the drafting stage may save the writer much work at the revising or proofreading stage later. In this regard, L2 learners may take every writing task as a language learning opportunity in which they can practice form-function-meaning mappings to make use of different functions of given reference resources.

The final perspective is that although reference resources can serve as a more competent interlocutor with whom the learner engages in meaning negotiations, they should not be considered tutors but rather informants at best (Gavioli, 2009). Unlike human teachers or tutors, reference resources are not intelligent by themselves but add to a user’s cognitive power only through her active participation in intellectual partnership with them. In other words, L2 writers should abandon the expectation that reference resources are tools that work for them by helping
them to complete the same tasks with less time and effort. Instead, reference resources are tools that one should work with in a cognitive division of labor. This approach means that writers still take full responsibility for finding solutions to their writing problems and making their own decisions, and consequently they still have to deal with uncertainty about the solutions they find through consultations. However, at the same time, writers need to realize that they are learning in the process and that the solutions arrived at through interactions with reference resources are more likely to be correct than ones that an unaided mind alone would produce. In this regard, L2 writers should also be aware that the use of reference tools works toward increasing their autonomy for writing independently.

Along with these facilitative perspectives, universally applicable strategies for reference tool use should be part of the general principles of reference resource consultation. I also developed a set of these general strategies that can be employed at each step of the problem-solving cycle (see Appendix H for detailed discussion of these strategies). Learners can be introduced to these general principles at the beginning of their training and reminded of them while taught about and practicing with individual reference resources.

Learner training can also include guidance on the benefits of using multiple (types) of reference resources. As evidenced in the findings on the query purposes and the resources consulted, concordancing cannot replace but only complements other types of reference resources, which have each their own unique functions useful for L2 writers’ problem solving. For this reason, while the primary emphasis should be on concordancing as it is a new and unfamiliar type of resource for most students, learner training may cover different types of reference resources, especially fast-evolving hybrid resources like online bilingual dictionaries (e.g., Naver). At the same time, while these resources are separately introduced and instructed, how to strategically use different resources in combination within a problem space and how to integrate information extracted from these different resources should constitute an important part of training. However, the mutual complementarity of different resources does not mean that more is always better (Frankenberg-Garcia, 2005). Writing teachers or consultants should choose a set of non-overlapping, complimentary resources that are appropriate for their students’ proficiency levels, interests, and task types.
Implications for reference tools

This study also provides some useful insights into reference resource and tool development. The majority of existing concordancing tools have been designed with linguistic inquiries by linguists in mind, and as a result their interfaces are often cramped with various sophisticated options. When used by learners, the complicated interfaces requiring the use of special search syntax and operators can be overwhelming and thus demotivating. Many researchers have therefore suggested that corpus tools used in language classrooms should be designed with pedagogical and learning purposes in mind and so with simple interfaces and only essential functionalities (Breyer, 2006; Perez Padres et al., 2012; Römer, 2006; C. Yoon, 2011). The present findings further suggest that concordaning tools should be even simpler and faster when used as reference resources for on-the-fly problem solving and that the corpora on which they run should be large rather than small. As evidenced in participants’ i-Conc use patterns, an L2 writer may often carry out a series of quick queries within a short span of time, engaging in robust meaning negotiations with the reference tool. Compared to when used for a DDL activity, where learners usually focus on interpreting concordance lines produced by one or two major queries, a concordancing tool used as a reference tool should thus have a more user-friendly interface and be fast in loading query results. In the meantime, for robust meaning negotiations where multiple input-feedback exchanges occur between the writer and the tool, large corpora are more suitable. As also shown in the results of the present study, an L2 writer would not be able to go very far in her negotiations with a small corpus like CSE, which often produces no or few instances of the target item. Therefore, as some participants mentioned, an ideal concordancer for immediate problem solving during writing would be one that combines the strengths of COCA and JTW: that is, the large POS-tagged corpus and relatively fast loading time of COCA and the learner-friendly interface and results display of JTW.

Another insight emerging from the present findings concerns an important type of inquiry that none of the resources in i-Conc could address satisfactorily. The present participants showed a high rate of consultation abandonment in problems of finding alternatives (both simple and upgrading alternatives). Many of these queries were intended to look for alternatives to multi-word items (e.g., point out and these days) whose searches are not allowed in most of the resources in i-Conc which only provide synonyms of single-word items. A certain portion of these queries, however, were not simply intended for synonyms in meaning but rather for
alternatives in function. For example, in HP48, Jinho performed the query \textit{[=in summary]} on COCA, with which he intended to find alternative expressions that can start off a summary or conclusion. This synonym search is not allowed in COCA in the first place as the synonym search syntax \textit{[=word]} can only be used with single words, but what Jinho wanted to do with this query was to find alternatives sharing the same communicative function. In her article laying out different challenges in using corpora to teach academic writing, Ädel (2010) pointed out as one of the most serious challenges the incapability of corpus tools to find alternatives having similar communicative functions (e.g., \textit{hedging} and \textit{expressing assumptions}), which can be realized in different surface forms. This expectation cannot be efficiently addressed in online dictionaries either. Thus, a new type of reference resources could usefully be developed (or existing resources extended) to provide multi-word synonym searches in meaning and function so that learners can be helped more effectively and efficiently with form-function-meaning mappings. One way of developing a tool with such capability would be to make use of existing resources. For example, Academic Phrasebank (\url{www.phrasebank.manchester.ac.uk}), is a website provided by Manchester University that features categorized collections of various phrases that can be used in different communicative functions (e.g., \textit{discussing findings}, and \textit{being critical}) in academic papers. This type of phrase banks could be extended and integrated into a larger corpus tool.

Lastly, the study results showed that the single interface of \textit{i-Conc} not only saved the participants the trouble of opening a window or browser tab for each resource, but the integrated interface also encouraged them to toggle between the different resources by having them physically next to one another in one place. Thus, when choosing a reference tool for their students, writing teachers or consultants may look for a tool or website (e.g., \url{www.lextutor.ca} and \url{garyshome.net/search}) that provides access to different types of resources and displays search results in the same interface, or they may wish to develop a simple program like \textit{i-Conc} with the help of a programmer.

7.3 Limitations of the study and directions for future research

Although this thesis has shed some new light on the under-researched areas of learner concordancing and reference resource use by using multiple methods, there were also some
limitations in its research providing an accurate and comprehensive picture of L2 writers’ reference resource consultation for their academic writing. Acknowledgement of these limitations here leads to various suggestions for future research, to address these limitations and to seek new directions, presented in the concluding part of this section and chapter.

7.3.1 Limitations of the study

First, many researchers have pointed out (Boulton, 2010a; Kennedy & Miceli, 2010; H. Yoon, 2005) that it is a long process for any learner to be able to engage effectively with any technology. This study examined the participants’ i-Conc use only for the duration of one writing assignment. Thus it was too short to observe L2 writers’ development of reference resource consultation literacy and changes it may bring to their writing behavior and performance over time. In particular, the three-hour tutoring sessions provided at the beginning of the study and the following few weeks of familiarization may not have been enough for the participants to sufficiently understand the features and functions of each resource of i-Conc. In fact, some of the participants (e.g., Jinho) were observed to still be not fully familiar with COCA until his last screen recording session and reported that they had not used the resources long enough to perceive tangible changes in their attitudes and approaches to L2 academic writing.

Another limitation is the great variability in the participants’ familiarity with the suite before and during the study and their writing assignments. Even though they were all Korean ESL graduate students at the same university, their duration of participation in the research and in their academic studies, and their familiarization period in particular varied widely. Their writing assignments in turn, while being all academic writing tasks, varied in type, the importance the participants imparted to them, and therefore their willingness to consult i-Conc while working on them. These cross-participant differences made it difficult to directly compare and contrast their i-Conc patterns. Next, although there were clear rationales guiding the selection of the present population of learners (e.g., examination of L1 use in and cultural orientations to L2 writing), having only Korean ESL students as participants limits the generalizability of the study results to broader L2 writer groups.
The study also had a few particular methodological limitations. One obvious limitation was the way screen recordings were conducted. As discussed in Chapter 3, screen recording itself is an innovative method that allows a researcher to observe the participant’s writing process unobtrusively. In the present study, however, the participants’ writing assignments were all major tasks for their academic degrees, which provided ecological validity for the research but also took up to several weeks to complete, so it was not possible to screen record the whole writing processes and analyze them. Instead, the study confined the observation of the writing processes to three one-hour screen recording sessions, and only the queries performed during these sessions were analyzed. Limiting the screen recording to only a few hours out of their entire composing processes this way may have made the participants particularly aware that they were being observed and it, in turn, may have caused them to feel pressured to consult i-Conc even when they did not feel a particular need to do so. Therefore, it is probable that the participants’ reference resource consultation patterns and behavior differed depending on whether their writing processes were screen recorded or not.

The perception of limited training and familiarization discussed above implies that some participants may have still been in the process of learning and practicing using the new resources while working on their writing assignments. Some portions of the queries may have been performed to test or explore the features and functions of certain resources rather than carried out with a concrete or routine purpose in mind. The study did not separate these queries from the ones that actually reflect the participants’ preferences and intentions.

Lastly, the categories developed for query purposes were not of equal granularity. Some categories (e.g., SC, simple confirmation) were so broad as to include many potentially subdividable purposes while others (e.g., POS, confirmation of the word class of an item) were perhaps too specific. For future studies on query purpose, the codes can be refined so that some of the minor categories are merged while some major one can be divided into finer-grained categories.

7.3.2 Suggestions for future studies

As others have often noted (Boulton, 2009; Chambers, 2007; Gaskell & Cobb, 2004), in order to popularize the use of learner concordancing along with other reference resources, more
empirical evidence drawn from wider ranges of populations and settings is needed. Therefore, future studies may look into L2 writers at different language proficiency levels in various settings including both academic and workplace contexts. As discussed in Section 7.1, overlapping but distinct contexts and purposes for writing (e.g., writing-to-learn-content, learning-to-write, writing-to-learn-language, or writing-to-mean contexts) may largely determine the types of writing, and the goals and needs of writers, and therefore affect the extent to which L2 writers perceive and act on the affordances offered by reference tools. Thus, different sets of consultation patterns, strategies, challenges, and evaluations can be expected to appear in studies following L2 writers’ reference tool uses, for example, in foreign language (not usually used in local communities) writing classrooms where writers are more motivated to attend to language forms in contrast to an expressive or creative writing course where writers engage in writing of a narrative or artistic nature. Combined, studies into diverse contexts and purposes for writing should provide insights into how to support L2 writers to exploit the full affordances of concordancing and other resources within different writing contexts and ultimately to develop on their own meaningful cognitive partnerships with given resources across contexts. Accordingly, one common thread in research in all these settings and contexts should be investigating how reference resources help to improve L2 learners’ confidence and autonomy as writers and ultimately serve as life-long intellectual partners.

Another avenue for future research is longitudinal studies that examine the extent to which the use of reference resources as a problem-solving tool while writing affects language acquisition and development. Specifically, L2 writers’ uses of reference tools can be traced over a year or longer to investigate the extent to which they affect writers’ attention to form, and types of problems they notice and address and how these change over time, focusing on how writing aided with reference tool use can be a site for internalizing existing knowledge and acquiring new knowledge on the target language. One promising point of departure in this line of research would be written error correction. This pedagogical practice has been a locus of intersection between studies of L2 writing and SLA (Ortega, 2012), and corpus tool use for written error correction has been one of the major themes in numerous concordancing studies (Chambers & O’Sullivan, 2004; Gaskell & Cobb, 2004; Gilmore, 2009; O’Sullivan & Chambers, 2006; Watson Todd, 2009). Future studies may expand on these previous studies by employing a longitudinal mixed methods design to provide rigorous evidence of links between writing,
reference resource consultation, and language development. This line of research would advance understanding about the interfaces between theories and research on L2 writing and SLA.

The present study only looked at the problems the participants chose to solve at the sentence level and below. Thus, another line of future research would be to investigate the potential of the combined use of multiple reference resources for improving various aspects of writing that go beyond word and sentence levels to discourse or rhetorical levels, as suggested in previous studies (Ädel, 2010; Charles, 2007). As more and more efforts will be made to create specialized corpora and corpus annotation designed to allow for searches by rhetorical function (along with meaning and surface form), and also searches for exemplar texts, future studies will need to consider how L2 writers utilize these resources for their writing. Along with research into new types of corpora and reference resources, efforts should continue to develop learner-friendly interfaces borrowing insights from related fields such as Human Computer Interaction.

The critical role that training can play in L2 learners’ reference consultation literacy development implies that learner training is another important research area to pursue. Given that systemic learner training in reference resource consultation is yet rarely practiced in classrooms, more practical studies are needed to explore cost-effective ways of training learners in basic skills and concepts required to use different reference tools, motivating them to notice unique affordances on their own, and to develop personal strategies for meeting their particular referencing needs. One promising research methodology is design-based research, in which teachers themselves, in collaboration with researchers if necessary, carry out multiple iterations of design, analysis, and implementation, working with multiple cohorts of students to develop contextually-sensitive learner training content and methodologies (Wang & Hannafin, 2005).

Lastly, the present study focused on participants’ cognitive processes involved in interactions with the reference suite from the relatively narrow angle of “cognitive tool”. Future studies may go beyond observing and analyzing the micro-levels of tool use by L2 writers to examine how the use of reference resources and other mediating tools “intersects with, and contributes to, broader social, cultural, historical, and economic trends” (Warschauer, 2005, p. 48).
REFERENCES


Bloch, J. (2007). *Technologies in the second language composition classroom*. Ann Arbor, MI:


John Benjamins.
Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 135-164). Boston, MA: De Gruyter Mouton.


Hubbard, P. (2004). Learner training for effective use of CALL. In S. Fotos & C. Browne (Eds.), *Perspectives on CALL for second language classrooms* (pp. 45-68). Mahwah, NJ:
Lawrence Erlbaum.


and learner evaluation. *Journal of Second Language Writing, 15*(1), 49-68.


and a wish list for future development. *Zeitschrift für Anglistik und Amerikanistik, 54*(2), 121-134.


Appendix A

Information Letter/Consent Form

Using online concordancing and other Internet-based reference resources as writing assistance: A mixed methods study of Korean ESL graduate students’ academic writing

Information Letter/Consent Form

Investigator: Choongil Yoon  Faculty Supervisor: Dr. Alister Cumming
Department of Curriculum, Teaching & Learning, OISE/University of Toronto

Purpose of the Study
This study investigates how Korean ESL (English as a second language) graduate students would independently use a suite of Internet-based reference resources (consisting of an online corpus and dictionaries) as a writing supporting tool to solve problems they encounter while writing in English and how the use of this reference tool would affect their attitudes and approaches to English writing. Through the inquiry, the study ultimately aims to gain insights into how learners can become more autonomous and confident in L2 writing.

Participants
Korean ESL students enrolled in University of Toronto graduate programs, who take a content course (no ESL or academic writing courses) that has at least one major writing assignment (10 pages or longer)

What Participants Will be Expected to do
This study involves observing and analyzing participants’ use of a suite of Internet-based reference resources while writing their course papers in English. The suite, developed for this thesis research with the help of Hwang-joong Kwon, a computer programmer and former colleague of the researcher, is a mini web browser that allows users to navigate and search in freely available online corpus and other language reference resources.

Data will be collected for about four months starting in November 2011.

If you agree to participate in the study, you will be asked to do the following:

1. Two surveys and three interviews
   Over the entire course of the study, you will fill out two surveys (30 minutes each) and three interviews (two one-hour and one 30-minute interviews) about your attitudes and approaches to English academic writing, use of reference resources, evaluation of the reference suite, etc.

2. Training
   Participants will receive three one-hour sessions of training in how to use the reference suite for supporting their English writing including basic concepts of each reference resource, and search strategies.
3. Screen recordings and stimulated recalls
   You will record your screen activities for 40 minutes while writing an English academic paper consulting the reference suite with a screen recording program installed on your computers. You will then conduct a stimulated recall, in which you watch the video file with the researcher and recall why you performed specific queries. There will be 4-5 cycles (depending on your schedules) of a screen recording and a stimulated recall for each of you over the entire course of the study.

4. All the interviews will be audio recorded and transcribed for analysis. The papers you write consulting the reference suite will be collected for analysis as well.

5. The reference suite, developed for the study, is hosted on a website and all the queries performed on the website will be recorded in the form of a query log and will be accessed by the researcher for analysis. The researcher may also access the query log even after the study is completed to check the participants’ continuous use of the suite.

The total amount of time required for participating in all the procedures mentioned above will be around 10 hours. Interviews, training sessions, stimulated recall sessions will be scheduled at participants’ convenience.

Participants’ Rights

- **To Confidentiality**
  - Data obtained from the study including surveys, query logs, audio and video files and written transcripts of interviews, and written work collected from participants will be treated with absolute confidentiality.
  - To further 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.
  - The data will be kept in a secure place. Only the researcher and the faculty supervisor mentioned above will have access to this information. Upon completion of this project, all data will be destroyed within one year.

- **To Ask Questions at Any Time**
  - You may ask questions about the research at any time. For questions, please contact the investigator at 647-863-2457 or choongil.yoon@utoronto.ca, or his supervisor Alister Cumming at 416-978-0276 or alister.cumming@utoronto.ca.

- **To Withdraw at Any Time**
  - Your participation in the study is absolutely voluntary. If, for any reason, at any time, you wish to stop participating in any parts of the procedures, refuse to answer any question, or withdraw from the study altogether, you may do so without any negative consequences. You may also require that your data be destroyed.

**Benefits**
You will be introduced to a variety of reference resources available on the Internet and functions and strategies you can use to solve language/writing problems that arise while writing in English. While doing so, you will have opportunities to reflect on your own attitudes and approaches to English academic writing and thereby develop a greater understanding of your
strengths and weaknesses in English writing. In appreciation of your participation, the researcher will also provide assistance in your writing of course papers that are not related to this study.

Risks
There are no risks or discomforts involved in this study other than your sharing with the researcher your usual approach to writing a course paper in English.

Signature
I certify that I have read and understood the above, that I have been given satisfactory answers to any questions about the research, and that I have been advised that I am free to withdraw my consent and to discontinue participation in the research at any time, without any prejudice or loss of benefits. I agree to allow Choongil Yoon to use findings, examples, and quotations from the study for purposes of publication, academic presentation, and teaching, provided that anonymity is maintained. For information about your rights as a research participant, please contact the Office of Research Ethics at ethics.review@utoronto.ca / 416-946-3273.

__________________________                                /      /
(print your name)                                          (date)

__________________________
(signature)
Appendix B

*i-**Conc Tutorial

Basic concordancing concepts and brief introduction of *i-Conc

1. Corpus and concordancing
- Corpus: a large, principled collection of naturally occurring texts (written or spoken) stored electronically. The BNC, COCA, etc.
- Concordancing: the process of using software (called concordancer) to search for all the occurrences of one word (or phrase) in a corpus.
- KWIC (Key Word In Context): the most widely used concordance format. All of the occurrences are presented with the node word/phrase (the one searched for) in the centre of the line, with 7-10 words presented at either side of the node word.

Figure 1. KWIC display of the phrase “to my utmost”

- Google search engine: Although not in the traditional sense, Google can be a concordancer with its search results showing the “query word/phrase” with its immediate context in the form of a “snippet”

Figure 2. Google search results of the query “to my utmost”
2. What corpus investigations tell us

- The open-choice principle (slots and fillers): traditional grammar
  e.g., Adjective + Noun, Adverb + Adjective
  A large range of choice is possible and the only restraint is grammaticalness
- The idiom principle
  A language user has available to him or her a large number of semi-pre-constructed phrases that constitute single choices, even though they might appear to be analyzable into segments.

*Strong tool vs Powerful tool; Strong opinion vs Powerful opinion*

1) Collocation
   The relationship between lexical item and other lexical items
   Words co-occur with certain words more often than others in actual use.
   Collocations are not absolute or deterministic, but are probabilistic, resulting from repeated combinations used and encountered by the speakers of any language (O’Keefe et al., 2007)
   e.g. make a decision (* do a decision), commit a crime (* make a crime)

2) Colligation
   The relationship between lexical item and a grammatical category
   e.g., “naked eye”, verb/adjective + preposition + the + naked eyes
   “see it coming” usually occurs in a negative or interrogative sentence.
   [put [app*] finger on], [anytime soon]

3) Semantic preference
   A word/phrase tends to co-occur with words/phrases belonging to a certain semantic set or sets.
   e.g. *Sheer* + 1) magnitude, weight: the sheer weight/size of …
        + 2) energy, force: the sheer force of …
        + 3) strong emotions: sheer joy/happiness
   *Naked eye* + words denoting ‘visibility’: detect, spot, appear, perceived, see/apparent, evident, obvious, etc.

4) Semantic prosody
   Words tend to occur predominantly in positive or negative environments (or pleasant/unpleasant)
   e.g. *cause* vs *provide*
   *cause*: accident, cancer, crisis, delay etc.
   *provide*: food, care, relief, support etc.
   *naked eye* has a semantic prosody of difficulty: too faint to be seen with the naked eye

   ● Semantic preference, syntactic patterning, meaning, genre, domain, and context all interact with one another.
   ● Deviation from these (lexico-grammatical) co-occurring patterns would produce various discourse effects (e.g. creative use of language, irony,
sarcasm, etc) when it’s intentional.

3. Corpus investigation and L2 writing
   - Corpus analyses have revealed the linguistic features of academic discourse and variations across academic genres and disciplines in terms of lexico-grammatical patterns and phraseology.
   - For L2 writers, concordancing can serve as a powerful reference tool to use to solve problems they encounter while writing in the target language.

4. *i-Conc*
   - *i-Conc* is a program that hosts a suite of Internet-based language reference resources that can be used for writing assistance.

   - It consists of four tabs (see Figure 3):
     1) **COCA**
        The Corpus of Contemporary American English  [http://corpus.byu.edu/coca/](http://corpus.byu.edu/coca/)
     2) **Google**
        a. Google Web Search ([www.google.com](http://www.google.com)),
        b. Google Scholar ([scholar.google.com](http://scholar.google.com)),
        c. Google Custom Search (different engines depending on the user’s specialty area)
     3) **Justtheword** ([www.just-the-word.com](http://www.just-the-word.com))
        a. Justtheword – combinations
        b. Justtheword – alternatives
     4) **Dictionaries**
        a. Bilingual Eng-Kor dictionary ([endic.naver.com](http://endic.naver.com))
        b. Monolingual Eng-Eng learner’s dictionary ([www.ldoceonline.com](http://www.ldoceonline.com))
        c. Thesaurus ([www.thesaurus.com](http://www.thesaurus.com))

![Figure 3](image)

   - Navigating across different tabs and resources are intuitive. The user just selects the tab and the resource he/she wants and type in the query word/phrase.

   - On each resource page, there are two search boxes: *i-Conc’s* own search box and another provided by each resource site. For example, Figure 3 shows the Google Custom Search page under the Google tab and there are two search boxes. The user may use either search box to perform queries, but only queries using *i-Conc’s* own
Search box (circled in red in Figure 3) will be stored in the query log in the server. **Study participants are encouraged to use i-Conc's search box so that queries they make can be recorded for analysis.**

- i-Conc requires registration (an email address and a password) for first time use and a login for subsequent uses.

- How to use each resource (basic functions and features)

  1) COCA
   a. The largest freely-available corpus of English (American English): the corpus contains more than 425 million words of text and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990-2011 and the corpus is also updated once or twice a year.

   b. Basic query syntax

   Table 1. Key query syntax

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Meaning</th>
<th>Examples</th>
<th>Sample matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>One &quot;slot&quot;</td>
<td>Make sure there is no space, or it will be interpreted as two consecutive words</td>
<td></td>
<td></td>
</tr>
<tr>
<td>word</td>
<td>One exact word</td>
<td>mysterious</td>
<td>mysterious</td>
</tr>
<tr>
<td>*</td>
<td>Any number of letters</td>
<td>re*ly</td>
<td>really, recently, relatively, etc</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>ous</em></td>
<td>house, serious, various, etc</td>
</tr>
<tr>
<td>?</td>
<td>One letter</td>
<td>s?n</td>
<td>son, sun, sin</td>
</tr>
<tr>
<td>[word]</td>
<td>All forms of the word</td>
<td>[do]</td>
<td>do, does, did, doing, done</td>
</tr>
<tr>
<td>[=word]</td>
<td>Synonyms</td>
<td>[=vicious]</td>
<td>mean, wild, violent, etc</td>
</tr>
<tr>
<td>[pos]</td>
<td>Part of speech</td>
<td>[vvg]</td>
<td>going, using, etc face (only as a verb)</td>
</tr>
<tr>
<td>word.[pos]</td>
<td>Word and part of speech</td>
<td>Face.[v*]</td>
<td></td>
</tr>
</tbody>
</table>

For the full list of query syntax, check [http://corpus.byu.edu/coca/help/syntax_e.asp](http://corpus.byu.edu/coca/help/syntax_e.asp)

  c. Major features

  [LIST]
  The LIST option gives a list of strings that match the query. A click on an item on the list will call up all the incidences of that item used in context.
Let’s say you are looking for the most typical or frequent adverbs that come before the adjective entrenched (that is, adverb collocates of entrenched). You may select LIST as above, and type \([r^*] \text{entrenched}\) (meaning all adverbs + entrenched). A click on the SEARCH button will return a list of all the “adverb + entrenched” combinations found in the corpus in the order of frequency as in Figure 4 below.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>CONTEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>FIRMLY ENTRENCHED</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>DEEPLY ENTRENCHED</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>SO ENTRENCHED</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>MORE ENTRENCHED</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>WELL ENTRENCHED</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>MOST ENTRENCHED</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>SOLIDLY ENTRENCHED</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>TOO ENTRENCHED</td>
</tr>
</tbody>
</table>

A click on one of these combinations will call up the incidences of the combination used in context (concordance lines). For example, if you click on firmly entrenched, the concordance lines that contain the phrase will come up as in Figure 6 below.
The KWIC option returns concordance lines that contain the query word/phrase in the KWIC (Key Word In Context) format. It is different from the LIST option in its display of concordance lines. The KWIC option does not provide all the incidences of the query word/phrase as the LIST option does but only 100-200 incidences (if the queried word/phrase occur more than 100 -200 times) with their co-texts color coded for different word classes (Figure 8). In this way, the user may notice syntactic environments a word or phrase typically occurs.

Using COCA within i-Conc (saving queries)

Unlike other reference resources in i-Conc, the COCA tab does not have i-Conc’s own search box. As COCA requires the user to choose several options for a query, which can be done only in its interface, queries cannot be
performed through i-Conc’s search box. Therefore, to store queries made on the COCA page, the user is encouraged to copy each query he/she types in COCA’s search box, paste it in the SAVE box, and press the SAVE button.

2) Google Search Engines

a. The Google tab in i-Conc features three Google search engines:

Table 2. Google search engines

<table>
<thead>
<tr>
<th>Engines</th>
<th>Scope of searches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Web Search</td>
<td>What we know as “Google”</td>
</tr>
<tr>
<td></td>
<td>Searches the entire Web</td>
</tr>
<tr>
<td></td>
<td>Provides rich data even for relatively infrequent words</td>
</tr>
<tr>
<td>Google Scholar</td>
<td>Searches academic publications across diverse disciplines</td>
</tr>
<tr>
<td></td>
<td>Useful when confining searches to academic registers</td>
</tr>
<tr>
<td>Google Custom Search</td>
<td>Customized to search only the web sites and linked pages the user designate in advance</td>
</tr>
<tr>
<td></td>
<td>Useful when further narrowing searches down to a certain domain or discipline</td>
</tr>
</tbody>
</table>

b. Google is arguably the most widely used search engine and most Internet users may already be familiar with its basic search functions and even have their own search strategies. Here some search tips are provided that can come in useful when using Google search engines as concordancers for language reference purposes.

Table 3.

<table>
<thead>
<tr>
<th>Search operators and features</th>
<th>Functions</th>
</tr>
</thead>
</table>

Phrase search: double quotes ("xxx")

By putting double quotes around a set of words, you are telling Google to consider the exact words in that exact order without any change. That is, Google returns the web pages that contain strings that exactly match the query words. Without the double quotes, Google returns web pages that contain any of the query words. This feature is useful especially because function words like ‘a’, ‘the’, and ‘of’ are usually ignored in Google searches when not put within double quotes.

Fill in the blanks: wild card (*)

If you include * within a query, it tells Google to try to treat the star as a placeholder for any unknown term(s) and then find the best matches. For example, the query [“a * deal of”] will return “a great deal of”, “a good deal of”, “a big deal of” etc. This feature is useful when you try to find a phrase you only partially know or collocates that co-occur with a specific word/phrase/pattern.

OR() operator

OR (all capital) within a query will return results that contain either one of words right before and after the OR operator. For example the query [“a great OR good deal”] will give you results that contain “a great deal” and “a good deal”. OR can be substituted by the symbol “|”. [“a great|good deal”] will return the same results.

allintext:

If you start your query with “allintext:”, Google restricts results to those containing all the query terms you specify in the text of the page. The query [allintext: “a great deal of”] will return web pages that contain “a great deal of” in their texts (not in their titles). This feature is particularly important when using Google as a concordancer because Google searches both titles and texts (body) of web pages for the query terms and web page titles are often elliptical and even ungrammatical.

Note: a. Square brackets [] denote an example of a query. For more detailed and further search features of Google, see http://www.googleguide.com/category/query-input

3) Justtheword

a. Using corpus linguistics methods, JustTheWord provides writing aid by helping writers choose the right words and suggesting improvements. Table 4 explains the function of each option.

Table 4. Options in JustTheWord

<table>
<thead>
<tr>
<th>option</th>
<th>function</th>
</tr>
</thead>
<tbody>
<tr>
<td>JTW combinations</td>
<td>A query (mostly a single word) made in this option will give you a detailed description of the company which that query word keeps in modern-day English. To help you find your way to the information you need, in the right-hand frame you'll find the part(s)-of-speech and the types of relation that the word is found in. If you're looking for the right adjective to modify a noun you've chosen, click on the ‘ADJ mod &lt;word&gt;’ link. If you want a verb with the noun as its object, follow the ‘V obj &lt;word&gt;’ link. Within many types of relation you'll find the uses of the word clustered into groups with a similar meaning.</td>
</tr>
</tbody>
</table>
JTW alternatives

A query (mostly two to three words) made in this option will give you an idea of how well these query words go together - a red bar indicates that the combination is unlikely (the longer, the unlikelier) - and some suggestions for improvements. For each word in the input, JustTheWord will try replacing it with a related word, and show you the strength of the combination in the usual way (green bars). The blue bars represent the similarity between the original word and its replacement.

4) Dictionaries

a. The Dictionaries tab has online versions of the traditional language dictionaries: a bilingual Korean-English (English-Korean) dictionary, a monolingual English learner’s dictionary and an English thesaurus. As these dictionaries have long been used by language learners for language reference purposes, and are in large part simpler to use than concordancers, how to use them will not be discussed here.

5) COCA vs Google Web Search

a. COCA and Google Web Search have their own merits and demerits as a concordancer. Table 5 compares these two.

<table>
<thead>
<tr>
<th>Table 5. COCA vs Google</th>
<th>Google Web Search</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pluses</strong></td>
<td></td>
</tr>
<tr>
<td>Genre-, grammar-based searches</td>
<td>Abundant data: much richer data than the biggest corpus</td>
</tr>
<tr>
<td>Targeted searches with various options</td>
<td>Quick and simple searches</td>
</tr>
<tr>
<td>Pattern searches</td>
<td>Can check whether even a long specific phrase (4-5 words) is used by others or not</td>
</tr>
<tr>
<td>KWIC format</td>
<td></td>
</tr>
<tr>
<td><strong>Minuses</strong></td>
<td></td>
</tr>
<tr>
<td>Limited amount of data compared to Google</td>
<td>Too much “noise”</td>
</tr>
<tr>
<td>No results for long phrase</td>
<td>May make infrequent expressions look frequent (a lot of examples of non standard usages, second or foreign language use) – alternatives: GS or GCE a</td>
</tr>
<tr>
<td>Complicated query syntax</td>
<td></td>
</tr>
</tbody>
</table>

Note: a. Search “Will I be ever able” and “Will I ever be able” in Google and Google Scholar.

Note: when you want to go back to the previous page, click anywhere on current page and press Backspace.
You may be required to type letters or digits from a distorted image, called CAPTCHA, sometimes while using Google search engines, which Google does to test if the user is a human.
5. Applications to Your Own English writing

- These resources can be good supporting tools for solving lexical and grammatical problems encountered while writing in English. Specific uses of these resources are varied and can also be developed by users themselves in ways that fit their learning/problem solving needs. The following is a classification of possible uses or purposes of reference resource consultation into three broad categories.

1) Verification
Verification refers to reference resource consultation for confirming whether a certain word/phrase or syntactic structure actually 1) is typically or frequently used in native speaker texts, or 2) is appropriate to use in a certain genre, and disciplinary writing, or 3) consultations for choosing the best from multiple alternatives.

2) Elicitation
Elicitation refers to consulting the resources 1) to find English equivalents to L1 words/phrases, 2) to find a typical and appropriate collocate or syntactic structure that co-occur with a certain word/phrase, 3) to see how different linguistic items can be combined in a sentence or paragraph, 4) to find alternatives to avoid the repetition of the same word/expression, or 5) to simply search for content that is related to the writing assignment at hand.

3) Improvement
Improvement refers to consulting the reference resources for finding 1) more accurate, effective and/or appropriate word/phrase, 2) more effective and appropriate syntactic constructions in the given context of writing than the one that the writer has written.

- General query strategies

- Focus on lexico-grammar: you should be aware that the reference tool can be more helpful to lexical and grammatical aspects of writing than to the aspects of discourse and organization.
- Trial and Error: to solve a certain language/writing problem may, more often than not, take more than a single query. You may have to refine your queries based on the results you get from the previous queries.
- Different reference resources have their own merits and demerits. COCA and Google are great tools but are not always the best or the most effective resources for finding what you are looking for. Some problems can be more efficiently solved by consulting more traditional resources like dictionaries. However, in many cases, it is better to use different resources in combination.

1) Verification
a. When in doubt about the accuracy of the query term, try to use the wild card (*).
b. To confirm the appropriateness of a certain phrase or word combination for academic register or genre, perform queries in Google Scholar or Custom Engine in the Google tab or select academic or other appropriate section of the corpus in COCA.
c. Practice
   i) “doesn’t influence on”
   ii) “will I ever be able”, “will I be ever able”

2) Elicitation
a. For an L2 equivalent, start with the bilingual dictionary, and look up the L2 word/phrase in other resources for its usage in context.
b. Query terms should not be too general or too specific.
c. When looking for a collocate, or syntactic pattern a certain word/phrase co-occur with, you may start with the learner’s dictionary, whose entries and subentries have been extracted from a large native speaker corpus. You can refine your search based on what you find in the dictionary and expand the search to other resources like Google and COCA.
d. Practice
   i) COCA: [j*] [assumption]; Justtheword combo: assumption (Adj + *assumption*)
   ii) COCA: [=basic] [assumption], a [=lack] of research
   iii) Google: develop interest

3) Improvement
a. At the lexical level, you may start with thesaurus and expand the search in Google and COCA.
b. You may use JustTheWord for alternatives.
c. Practice

- General strategies for results interpretation
  1) Beware of false positives (“influence on”)
  2) Pay close attention to whether the words you want to borrow are used in the search results exactly as you intend to use them in your text
  3) Check frequency info or the number of tokens
  4) Check the sources (genre, domain, URL, etc)
  5) Take care not to plagiarize what you find on the Web or COCA. You may need to contextualize what you find from the reference tool for your own writing.
  6) When using Google and dealing with queries that generate ample results, start your observation from page 5 or later. Google tends to return webpages from youtube, dictionaries and glossaries, song lyrics sites, etc. first, which might not be typical incidences of the searched terms.

6. How to install and use the screen recorder
   - The study participants are asked to screen-record the process of their writing for
about 40 minutes at a time (4-5 times) over the entire course of completing a course paper. The following is the instructions for installing and using a screen recording program (BB Flashback Express, a freeware program) on your computer.

1) Installation
   a. Go to [http://www.bbsoftware.co.uk/BBFlashBackExpress/Home.aspx](http://www.bbsoftware.co.uk/BBFlashBackExpress/Home.aspx) and press the Download button.
   b. Follow the instructions given on the download page: type in your email address and click the “Get Download Link.”
   c. Open the email they sent you and click the link provided in the message.
   d. Once the setup file (bbfbex3.exe) has been downloaded, install it by double-clicking it. Follow the installation instructions.
   e. When everything goes okay, two programs will be installed: BB Flashback Player and BB Flashback Recorder.
   f. When you launch the Recorder for the first time, a trial-period reminder window will pop up that asks you to register. Press the Register Now button.
   g. Fill out the required fields and press the Submit button.
   h. Go to your email account and open the email they sent you. Click the activation link. Now you can use the software for free with no period limitation.

2) Recording a video
   a. Open BB Flashback Recorder and click “Record a New Movie.”
   b. Select “Full Screen” (You can choose to record a specific window or region too, but for this study, full screen recordings are needed). Do not check “Record Sound” and “Record Webcam.” Check the “Minimize to Tray while Recording” option.
   c. Push the red “Record” button.
   d. While recording is in progress, a red light will be blinking in the small BB FB icon in the task tray in the bottom right-hand corner of the screen.
   e. When you think you have recorded enough, right click the BB FR icon in the tray and select “Stop” on the context menu.
   f. Press “Save” and choose a folder you want to save the file in.
Appendix C

Questionnaires

Initial Survey

1. Demographic information
1a Age _____ years old
1b Gender Female ___ Male ___
1c Major Humanities ___ Social Sciences ___ Other (please specify) ___
1d Degree pursued PhD, EdD ___ MA, MEd ___
1e Years of education in Korea and last degree earned there ____ ____
1f Years studying in Canada and other English speaking countries ___ years ___ months
1g English writing test scores: writing section of TOEFL ___, GRE writing ___, other ___

2. English academic writing

2a Have you ever taken English writing classes/courses? If you have, please explain when and what type of classes/courses you took.

__________________________________________________________________________

2b What is your overall attitude toward writing course papers in English?

Choose one from 1 (hate it) to 5 (enjoy it a lot)
1 2 3 4 5

2b How would you rate your overall confidence in English academic writing?

Choose from 1 (not confident at all) to 5 (very confident)
1 2 3 4 5

2c How difficult do you find each of the following aspects of academic writing in English? Choose from 1 (not difficult at all) to 5 (very difficult)

Vocabulary/Usage
1 2 3 4 5
Grammar/Sentence construction
1 2 3 4 5
Academic style/register
1 2 3 4 5
Rhetoric/organization
1 2 3 4 5
Other (please specify) ___________________
1 2 3 4 5
2d What would you like to improve most among the following areas of your English academic writing? Choose one.

- Vocabulary/usage ___
- Grammar/Sentence construction ___
- Academic style/register ___
- Rhetoric/organization ___
- Other (please specify) ___

3. Reference resource use

3a Do you use reference resources (e.g. dictionary) while writing in English?
- Yes ___
- No ___

3b If yes, what kind of reference resources do you usually consult? Check all that apply.
- Bilingual dictionary (Korean-English) Online ___ Paper ___
- Bilingual dictionary (English-Korean) Online ___ Paper ___
- Monolingual dictionary (English-English) Online ___ Paper ___
- Electronic handheld dictionary ___
- Grammar books ___
- Thesaurus ___
- Collocations dictionary ___
- Google or other Internet search engine ___
- Corpus tool ___
- Other (please specify) ___

3c Please check three that you use most
- Bilingual dictionary (Korean-English) Online ___ Paper ___
- Bilingual dictionary (English-Korean) Online ___ Paper ___
- Monolingual dictionary (English-English) Online ___ Paper ___
- Electronic handheld dictionary ___
- Grammar books ___
- Thesaurus ___
- Collocations dictionary ___
- Google or other Internet search engine ___
- Corpus tool ___
- Other (please specify) ___

3d For what purposes do you usually consult the reference resources? Check all that apply.
To see meanings of a word or phrase
To see usage examples of a word or phrase
To find an English equivalent to a Korean word or phrase
To check grammar related to a specific word or phrase
To find collocates that go with a word or phrase (e.g. make a decision, take medicine)
To confirm whether a word or phrase I am trying to use is accurate or actually used by native speakers
To check spelling
Other (please specify)

3e Which reference resource do you consult most for each of the following purpose?
For each purpose that applies, choose just one resource and write down the number assigned to each resource.

| 1 Bilingual dictionary (Korean-English) | 2 Bilingual dictionary (English-Korean) |
| 3 Monolingual dictionary (English-English) | 4 Grammar books |
| 5 Thesaurus | 6 Collocations dictionary |
| 7 Google or other Internet search engine | 8 Corpus tool |
| 9 Other (please specify) | |

To see meanings of a word or phrase
To see usage examples of a word or phrase
To find an English equivalent to a Korean word or phrase
To check grammar related to a specific word or phrase
To find collocates that go with a word or phrase (e.g. make a decision, take medicine)
To confirm whether a word or phrase I am trying to use is accurate or actually used by native speakers
To check spelling
Other (please specify)

Evaluation Survey

Here are some open-ended questions about your experiences of consulting the reference suite while writing in English over the last few months. Please take a moment to reflect on each of the following questions and share your thoughts. You can answer either in Korean or in English.

1 What are the biggest advantages of using concordancing (COCA, Google search engines) and other Internet-based resources in the reference suite while writing your course papers?
2 What are the challenges and difficulties you experience when using the reference suite?

Have there been any changes in your English writing in terms of attitude, confidence, writing process, and aspects to focus on etc since consulting concordancing (COCA and Google search engines) and other resources? If so, can you explain those changes?

4 Overall, how would you evaluate the reference suite as a writing supporting tool? Would you continue to use it for your writing assignments in the future?
Appendix D

Interview Guide

Date ______________________  Participant ID ________________

Introduction

☐ Introduce yourself
☐ Discuss the purpose of the interview
☐ Provide structure of the interview (audio recording, taking notes)
☐ Ask if they have any questions
☐ Test audio recording equipment

Questions for Interview 1 (Phase 1)

Attitudes and approaches to writing course papers in English

1. What is your overall attitude toward writing course papers in English? How confident are you about academic writing in English?

2. Tell me about difficulties you usually experience while writing academic papers in English?

3. What is the most important goal or standard you set if any when engaging in an English writing task? (e.g. producing an error-free text, or clear and coherent text) Do you usually set a specific goal for a writing assignment?

4. What is your typical process of writing a course paper like? (e.g., planning, drafting, revising) How long does it usually take to write a course paper? Are you likely to review what you have already written before continuing to write new material in the paper? Do you postpone editing your writing until you have gotten your ideas down?

5. What is your typical editing and proofreading process like when writing in English? What do you usually focus on? (e.g. vocabulary, grammar, register etc.)

Getting writing support

6. Where do you get support when you need help for your English writing?

7. What kind of support in what areas of writing do you get from different sources?

8. Is the support you get as sufficient and available as you would like it to be? In what aspects of writing do you need more support?

Corpus and reference resource use while writing
9. Do you consult reference resources (e.g., dictionaries) while writing? What kinds of resources do you consult for what purposes?

10. Do you use Google or other Internet search engines while writing? For what purposes?
    Do you usually find what you look for on the Internet? How would you rate your skills in Internet search?

11. How familiar are you with corpus or concordancing? Have you ever used corpus tools for language learning or writing?

12. How would you rate your overall computer skills (word processing, Internet search etc.)?

13. What is your overall attitude toward using computer and ICT while writing?

Questions for Interview 2 (Phase 2)

1. How would you evaluate concordancing as writing assistance for your academic writing?
   1.1 What is the single biggest advantage of using concordancing for your English academic writing?
   1.2 How does it help you solve problems while writing and/or improve your academic writing skills?
   1.3 Do you think concordancing and other resources in the suite complement one another? If so, how?
   1.4 What is the single biggest disadvantage of consulting concordancing for your English writing?
   1.5 What are the suite’s limitations as a writing supporting tool? What are the difficulties you have experienced using the tool?
   1.6 Taking together all that has been said above, what would be your single biggest reason for using (or not using) the reference resources or any specific resource while writing in English?

2. Have there been any changes in your attitude, and approaches to and process of writing course papers while using the reference suite?

3. Does the use of the suite affect your confidence in English writing and if so, how?

4. How can the suite be improved in terms of interface and features to be more helpful?

5. Are you willing to continue to use the reference suite in the future? Why and why not?

6. Is there anything else you would like to add or share besides what we have already talked about?
Concluding Statement
☐ Thank them for their participation
☐ Record any observations, feelings, thoughts and/or reactions about the interview
Appendix E

Stimulated Recall Guide

☐ Explain the procedure
☐ Assure the participant that he/she doesn’t have to be shy, self-conscious about sharing their thought processes
☐ Ask if the participant has any question

Instructions for participants:
We are going to watch a video you recorded of writing your paper. I am interested in what prompted you to consult the reference suite while you were writing the paper. Specifically I’d like you to tell me 1) what specific purposes you had in mind for a specific query you performed with the tool, 2) why you chose a specific resource or combination of resources to achieve the purpose, 3) how you came up with the specific query formulation, 4) how you made a decision on what to take from the results and how to apply it to your own writing.

I already watched the video and marked where you consulted i-Conc. As we watch the video, I will fast forward the video and pause at those marked parts to ask you a few questions about the queries you performed. If you want to tell me something about what you were thinking while writing you can pause the video at any time as we watch the video. This session will be audio recorded.

Questions to be asked:
Actual wording of the questions may vary depending on participants’ specific queries. The following are some possible examples of these questions.

Here, you looked up xxxxx. What did you perform this query for? (why did you look it up?)
Why did you choose to look it up in this specific resource?
The search results show there are a couple of options you can choose from. Why did you choose this specific one?
You discontinued this query here and started a new one. Can you explain why you gave up the search?
...

☐ Thank them for their participation
☐ Ask them if they have anything else to add or share regarding their experience of consulting the suite.
Appendix F

An example of a problem space

• Initial state: [...] more devastating images and videos of youth culture and life have been produced and shared online

<table>
<thead>
<tr>
<th>Problem: devastating images</th>
<th>Problem type: Confirmatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query</td>
<td>Resource</td>
</tr>
<tr>
<td>&quot;devastating images&quot;</td>
<td>Google</td>
</tr>
<tr>
<td>devastating images</td>
<td>COCA</td>
</tr>
<tr>
<td>[=devastating]</td>
<td>COCA</td>
</tr>
<tr>
<td>disturbing image</td>
<td>COCA</td>
</tr>
</tbody>
</table>

• Goal state: [...] more disturbing images and videos of youth culture and life have been produced and shared online

This is an example taken from Yumee’s data. “Initial state” is what Yumee initially wrote. She wanted to check whether the phrase “devastating images” delivers her intended meaning. She intended to say that these images make people uncomfortable, or irritated. So she performed the query “devastating images” on Google to check whether people use the phrase in the same way she used it. However, Google returned a great number of Web pages on natural disasters and major accidents. She wanted to make sure if the phrase is usually used in this context and performed the same query on COCA. There were only 2 tokens, which were also about natural disasters. Now she tried to find an alternative adjective to “devastating” and ran a query on COCA for synonyms of “devastating”. She chose from the synonym list the word “disturbing” and ran yet another query typing in “disturbing image” on COCA to confirm whether there were any instances in the corpus. She verified that it’s a frequently used phrase and also carried her intended meaning. Now she replaced “devastating” with “disturbing” in the sentence. This whole process from the initial state to the goal state is one problem space consisting of four queries. Problem type here is confirmatory and each query within the problem space has its own purposes, and function, and belongs to a linguistic category as seen in the last two right columns of the table above.
## Appendix G

### Possible Code Sets and the Solution Path Each Code Set Represents

<table>
<thead>
<tr>
<th>Group</th>
<th>Code set</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>S, C, Blank</td>
<td>The participant was satisfied with the solution she found from <em>i-Conc</em> consultation results and applied it to her writing. It resulted in a correct text formulation or revision.</td>
</tr>
<tr>
<td>B</td>
<td>D, C, Blank</td>
<td>The participant was not satisfied with the consultation results as the consultation did not return a clear or immediate solution to the problem. However, the participant nonetheless retrieved what she believed may be a solution to the problem and applied it to the writing, and it resulted in a correct text formulation or revision.</td>
</tr>
<tr>
<td>C</td>
<td>S, I, Blank</td>
<td>The participant was satisfied with the solution she found from <em>i-Conc</em> consultation results and applied it to her writing, but it resulted in an incorrect text formulation or revision.</td>
</tr>
<tr>
<td>D</td>
<td>D, I, Blank</td>
<td>The participant was not satisfied with the consultation results as the consultation did not return a clear or immediate solution to the problem. However, the participant nonetheless retrieved what she believed may be a solution to the problem and applied it to the writing, but it resulted in an incorrect text formulation or revision.</td>
</tr>
<tr>
<td>E</td>
<td>D, C, A</td>
<td>The participant was dissatisfied with the results of the first query or queries, abandoned the consultation, and attempted to solve the problem with no input from <em>i-Conc</em>. It resulted in a correct text formulation or revision.</td>
</tr>
<tr>
<td>F</td>
<td>D, I, A</td>
<td>The participant was dissatisfied with the results of the first query or queries, abandoned the consultation, and attempted to solve the problem with no input from <em>i-Conc</em>. However, it resulted in an incorrect text formulation or revision.</td>
</tr>
<tr>
<td>G</td>
<td>D, N/A, A</td>
<td>The participant was dissatisfied with the results of the first query or queries, abandoned the consultation, and also gave up the text formulation/verification or revision of the given content altogether. Therefore, there was no part in the participant’s paper which can be linked to the given <em>i-Conc</em> consultation and the problem solving could not be rated “correct” or “incorrect.”</td>
</tr>
</tbody>
</table>
Appendix H
Principles of Effective Reference Resource Consultation

The findings of this study show that effective use of concordancing and reference resources in general requires not only the acquisition of skills and strategies needed to use specific resources but also certain facilitative attitudes and perspectives on the use of reference resources. While technological advances have come with new types of reference resources and also changed the scope and functions of traditional tools, understanding these general principles will become more and more important for educators, learners, and writers. Some attempts have already been made to formulate basic principles of how to use corpus references effectively (e.g., the “apprenticeship” approach to corpus consultation for writing developed by Kennedy and Miceli, 2001, 2010). Chambers and O’Sullivan (2004) even argued that the concept of literacy generally should now include “the ability to select, evaluate and use the electronic tools and resources appropriate for the activity which is being undertaken” (p. 158). Focusing on the learner’s corpus use, O’Sullivan (2007) further proposed the promotion of “corpus consultation literacy” as a new but necessary form of literacy that language learners should be taught.

In what follows, in an attempt to answer to this call by O’Sullivan, I suggest a set of principles of reference resources consultation as problem-solving assistance while writing that can be commonly applied to different concordancing tools and other reference resources. This is also my attempt at theory building driven by practical application and interventions, one of the major principles of design-based research. The suggestions I make here overlap to some degree with what Kennedy and Miceli (2001, 2010) proposed for their apprenticeship approach, but I expand on it by incorporating insights drawn from the present study. The discussion is organized broadly in two parts. The perspectives conducive to effective reference resource use have been discussed in Section 7.2.2.3, so I present here only specific strategies of resource consultation for different steps of the consultation sequence.

Strategies for effective reference resource consultation

There are specific strategies that L2 writers can employ at the different steps of a consultation cycle (see Figure 4.2 for an example). Some of these steps are already documented
in Kennedy and Miceli’s studies and others arise from the findings of the present study (e.g., see Section 4.4 strategies and pitfalls in i-Conc consultation).

**Problem definition and representation**

Unless the writer consults resources to see what comes up in the query results with an intention of discovering syntactic patterns and vocabulary by chance, it is always better to start a consultation with a clearly defined and properly represented problem. That is, an efficient and effective consultation starts by defining what the problem is about and representing it in a specific form of question. This step of the consultation cycle is crucial as it guides how to devise a search strategy, which includes choosing a particular resource to consult, selecting a specific option to use within the given resource, and formulating a query term to enter.

Specifically, problem definition here refers to identifying the content of the problem. At the beginning of the present study, one of the mistakes the participants often made was to start their i-Conc consultations without clearly defining the purpose of the query. Defining the content of the problem is important as it determines what to look for in the query results. For example, Jinho was often observed to enter a single word he found problematic in the search field of a resource that he haphazardly chose, only to be puzzled about what to look for from the concordance lines. Rather, upon noticing a problem from her text, the writer may ask herself what exactly she wants to know: for example, whether to check syntactic patterns in which a word usually occurs or find collocates that typically occur with the word.

Once the content (e.g., general usage or collocation) of the problem is decided, the user now has to represent the problem into the form of a specific question, which can take one of two forms: *whether* question (closed form) and *what/how* question (open form). In this study, these two forms of questions, mostly formulated in the writer’s head, were externalized into two corresponding types of queries: verification and elicitation. As shown in the results, these two types have their own merits and demerits. A verification query can in general be better used when the writer is to some degree confident about the accuracy of the item in question and wants to get a quick confirmation, but can often bring the writer to a dead end when it fails to give the writer a successful confirmation. An elicitation query, in turn, is usually used to elicit linguistic items the writer cannot access on her own, but requires a strategic formulation of the query term. Therefore, the writer is advised to choose the more appropriate query type according
to her intention and constraints for a given problem. One thing to avoid is habitually representing problems in only one form of question. As shown by some participants in the study, by becoming fixated on only one type of query, the writer can miss out on opportunities to reach better solutions or the same solutions more efficiently.

**Query strategies**

Once the writer has a specific question in mind, the next step is to choose a resource to consult. Different resources have unique functions or purposes they are best suited for, though there is also considerable overlap. Concordancing tools are in general more effective at showing lexico-grammatical patterns and frequency information whereas more traditional resources can be better used when checking meanings, L2 equivalents, and synonyms. However, it should be kept in mind that each type of resources are not always better for these respective sets of purposes mentioned above. Instead of rigidly linking the resources to specific functions, users can be flexible in choosing resources depending on the query. For example, a simple elicitation query for typical/frequent syntactic patterns or phraseologies of a word can be carried out more quickly and efficiently on monolingual learner’s dictionaries, most of which present the relevant information drawn from large corpora. Likewise, meanings of phrases, which are often missing in dictionaries, can be explored by perusing concordance lines and some mediated concordancers like COCA provide a powerful synonym search function.

Another important principle in choosing resources is that when solving a specific problem, the writer does not have to choose just one resource. As was well demonstrated in the study results, strategically using multiple resources in combination can benefit L2 writers in a number of ways. First, it can provide corroborations for decisions the writer makes and therefore greater confidence in the solution she has found. Second, different resources complement each other, making it possible for the writer to reach solutions that would not be attained through a resource alone. For example, to find L2 equivalents, the writer can run a query on a bilingual dictionary to get different alternatives and then run a query of each alternative on a concordancer to check through the concordance which one best corresponds to the intended meaning.

Determining a query term to enter in the search field is also an important factor for a successful reference resource consultation. It is particularly so with concordancers such as
COCA and Google search engines because dictionary-type resources and JTW are usually used for single-word queries and therefore pose less difficulty in terms of query term formulation. One major challenge L2 writers may encounter is deciding how general or specific a query should be. Overly general queries (e.g., single words, frequently used word sequences, etc.) on a large corpus will generate too many concordance lines. Unless the user intends to check the frequency of the queried term in a given corpus or simply to observe what comes up in the results, it would be too much work for the writer to go through to find what she sets out to look for regarding a given linguistic item. For example, a query of a preposition or polysemous word in COCA or Google would not be in most cases a practical search. Thus, it is advisable that the writer should think first about what kind of results her query will produce, and how many irrelevant sentences will be returned along with the relevant ones, given the generality of the query term. Conversely, overly specific queries (e.g., a long word sequence, the use of only one inflectional form of a word or lemma, and semantically narrow words, etc.) often do not return enough matches to provide evidence regarding the queried item. In that situation, the writer may make her query more general by using a wildcard, “OR” operator, or hypernym, or may perform further queries using different lemmas or different words from the same semantic groups.

The last point is especially important in searches on collocations and formulaic language. One major trap the present participants were often observed to fall into was formulating a query term that does not allow for intervening words. For example, the participants often forgot that collocates do not necessarily occur right before or after the node word. Thus learners should be reminded that when formulating a query of a word string, a wildcard or placeholder can be used for possible intervening words or a specific number can be set for the span within which a collocate comes.

**Evaluation and application strategies**

In a reference resource consultation cycle, choosing the right resources and formulating the appropriate query term is half the battle. Then the other half is to evaluate the search results and apply the solution, if found from the results, to the text being constructed or revised.

One overarching principle in evaluating and applying the results is that the user should not observe the target item in isolation but pay as much attention to the co-text accompanying it. In light of this, one of the most common pitfalls learners tended to fall into was choosing the
most frequent item among multiple alternatives without confirming whether it matched the intended meaning or if its meaning differed from less frequent ones. This pitfall occurred often with mediated concordancers running on a POS-tagged corpus such as COCA and JTW. A query in these resources returns the results in the form of a list displaying word strings that meet the query condition in order of frequency. A click on each item on the list, in turn, retrieves concordance lines for that item. The participants often chose the top choice on the list and applied it to their writing without clicking into actual concordance lines of that item and thereby checking whether it was commonly used in the intended context, or by examining still other items on the list, one of which would have been what they were actually looking for. Even when using unmediated concordancers (i.e., Google search engines), the participants often checked the presence of the bold-faced target item in the results, while paying only minimum attention to the contexts it was used in. Under time pressure, one may often be tempted to do so, and therefore it is advisable to make a habit of observing the target item together with the context it occurs in.

A case in point is searching on prepositions, which are one of the most commonly searched-on linguistic items by the participants in concordancing studies. The writer should pay particular attention to whether the preposition in question is a part of a construction that comes before or after it in the concordance line. For example, a COCA search for prepositional collocates of a particular verb will return a list of combinations of “verb + preposition” (e.g., go on) in order of frequency. The writer should remember that a preposition that follows a verb can be either a part of the phrasal verb (e.g., we can’t go on like this anymore) or a part of a prepositional phrase that comes after (e.g., I will go on July 15th). In the latter case, the preposition is not a collocate of the phrasal verb that means continue.

Lack of rigor in observing the co-text can adversely affect the application of solutions found to one’s writing. As presented in Section 4.4, even when the present participants found a correct answer (e.g., an appropriate collocate), they sometimes produced an ungrammatical text by not paying close attention to the syntactic pattern in which it occurred. For example, even though the target item was consistently used with the indefinite article or in the plural, the participants missed the article and used it in the singular in their text.

Finally, when interpreting the query results, the writer should take into account the limitations of the resource consulted. Although each of the multiple resources used in the study
has its own strengths as a reference tool, it also comes with limitations that can complicate interpretation and evaluation of the search results. To discuss only a few of these limitations, COCA cannot display query results of a polyseymous word by its different senses, so the user should peruse the concordance lines to distinguish the different senses or refine the query in a way that would return only the matches relevant to a specific sense of the word. Google, due to its inclusiveness of language varieties, can give false confirmations on ungrammatical or stylistically inappropriate items. Lastly, when consulting concordancers running on a relatively small corpus, such as JTW and CSE, the user should be careful not to equate non-existence of instances with non-existence of the language phenomenon. As non-existence of the instances of a linguistic item is often caused by the small size of the corpus, the user may try the same query in a different resource.

Figure 1 summarizes the possible strategies that L2 writers can use at the different steps of the reference resource consultation cycle. As these strategies have been in large part drawn from the present study alone, they are far from being comprehensive and or sufficiently detailed to cover all reference consultation situations, but rather should be regarded as some basic rules further studies and educational practices can build on.

| Strategies for problem definition, and representation | • Define the content of the problem  
• Represent the problem as a specific question  
• Use an appropriate form of question (either open or closed form) that serves better the purpose of inquiry  
• Try not to stick to only one form |
|---|---|
| Querying strategies | • Be flexible and avoid rigidly linking specific functions to specific resources  
• Remember that different resources can be mutually complementary, and try to use them in combination while solving a problem  
• When formulating/refining a query, think about what the query results will be like and adjust the generality/specificity of the query term  
• When searching on a word sequence, allow for intervening words |
| Evaluation, application strategies | • Pay close attention to the co-text surrounding the searched-for item in the concordance lines  
- when choosing the most frequent alternative, make sure that it matches the intended meaning by checking its concordance  
- For a search on a preposition, examine the concordance lines whether it is linked to a construction to the right or left.  
- When applying the solution found to writing, make sure that the item is grammatically applied in terms of article/determiner, noun ending, |
tense, etc.

- When evaluating the query results, take into account the limitations of the resource consulted.

*Figure 1.* Strategies for different steps of reference resource consultation.