2019

Developing and evaluating corpus-based feedback

Mo Chen
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Bilingual, Multilingual, and Multicultural Education Commons, and the English Language and Literature Commons

Recommended Citation
Chen, Mo, "Developing and evaluating corpus-based feedback" (2019). Graduate Theses and Dissertations. 17162.
https://lib.dr.iastate.edu/etd/17162

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Developing and evaluating corpus-based feedback

by

Mo Chen

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Applied Linguistics and Technology

Program of Study Committee:
Carol Chapelle, Major Professor
Gary Ockey
Gulbahar Beckett
Evgeny Chukharev-Hudilainen
Connie Hargrave

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University
Ames, Iowa
2019

Copyright © Mo Chen, 2019. All rights reserved.
# TABLE OF CONTENTS

LIST OF FIGURES .................................................................................................................. vi

LIST OF TABLES ....................................................................................................................... viii

NOMENCLATURE ..................................................................................................................... x

ACKNOWLEDGMENTS .............................................................................................................. xi

ABSTRACT ................................................................................................................................. xiii

CHAPTER 1. INTRODUCTION ................................................................................................. 1
  1.1. Statement of Problems ................................................................................................. 1
  1.2. Goals of the Dissertation ......................................................................................... 3
  1.3. Significance of the Study .......................................................................................... 4
  1.4. Outline of the Dissertation ...................................................................................... 5

CHAPTER 2. LITERATURE REVIEW ....................................................................................... 6
  2.1. Corrective Feedback and Issues in Corrective Feedback Research ......................... 6
      2.1.1. Effects of Different Types of Corrective Feedback on Language Learning .......... 7
      2.1.2. Effects of Corrective Feedback on Correcting Syntactic Errors ....................... 12
  2.2. Learner Corpora and Corpus-based Tools for Error Correction in L2 Writing Classrooms ........................................................................................................... 15
      2.2.1. Direct and Indirect Use of Learner Corpus Data in CALL Tool Development in L2 Writing Classrooms .................................................................................... 17
      2.2.2. The Evaluation of Corpus-based Pedagogical Materials and Tools in CALL .... 20
  2.3. Theoretical Foundations for the Development of Corpus-based Corrective Feedback ...................................................................................................................... 25
      2.3.1. The Interactionist Approach and Corrective Feedback ........................................ 25
      2.3.2. Skill-Learning Theory and Corrective Feedback ................................................ 28
      2.3.3. Implications of SLA Theories for Pedagogical Design of Corrective Feedback ..... 29
  2.4. Theoretical Framework for the Evaluation of CALL Corrective Feedback ............... 32
      2.4.1. The Logic Model of the Theory of Action and the Development of the Logical Framework for Evaluating CALL Corrective Feedback ............................................. 33
      2.4.2. Chapelle’s Criteria for CALL Task Appropriateness and the Development of the Logical Framework for Evaluating CALL Corrective Feedback ............................ 35
      2.4.3. Research Questions ............................................................................................. 39
  2.5. Chapter Summary ......................................................................................................... 40
CHAPTER 3. METHODOLOGY

3.1. Research Design............................................................................................................. 42
3.2. Participants and Sampling............................................................................................ 44
3.3 Teaching Context............................................................................................................. 46
3.3. The Focused Error Types of this Study and Their Definitions................................. 47
3.3 Materials ....................................................................................................................... 49
  3.3.1. My Feedback Website ......................................................................................... 50
  3.3.2. Error Correction Exercises in My Feedback ...................................................... 57
  3.3.3. Error Awareness Sheet ....................................................................................... 59
  3.3.4. Pre- and Post-tests ............................................................................................ 59
  3.3.5. Background Survey Questions .......................................................................... 60
  3.3.6. Likert-scale Questionnaire ................................................................................. 60
  3.3.7. Instructions for the Think-aloud Protocols ......................................................... 61
  3.3.8. Semi-structured Interview Questions .................................................................. 62
3.4. Procedure ..................................................................................................................... 63
  3.4.1. In-class Activities ............................................................................................... 64
  3.4.2. Out-of-class Research ....................................................................................... 65
3.5. Data Analysis .............................................................................................................. 66
3.6. Chapter Summary ....................................................................................................... 84

CHAPTER 4. RESULTS ON THE LANGUAGE LEARNING POTENTIAL OF CALL CORRECTIVE FEEDBACK

4.1. Effects of Corpus-based and Coded Feedback on Error Noticing (Q1) .................... 85
  4.1.1. Quantitative Data Results ................................................................................ 86
  4.1.2. Qualitative Data Results ................................................................................ 89
  4.1.3. Triangulation of Quantitative and Qualitative Results ..................................... 98
4.2. The Effects of Coded and Corpus-based Feedback on Error Correction (Q2) ........ 100
  4.2.1. Quantitative Data Results .............................................................................. 100
  4.2.2. Qualitative Data Results .............................................................................. 104
  4.2.2. Triangulation of Quantitative and Qualitative Results .................................... 114
4.3 The Effects of Coded and Corpus-based Feedback on Knowledge Application (Q3) ........................................................................................................... 115
  4.3.1 Quantitative Data Analysis .............................................................................. 116
  4.3.2. Qualitative Data Analysis .............................................................................. 121
  4.3.3. Triangulation of Quantitative and Qualitative Results .................................... 130
4.4. Chapter Summary...........................................................................................................131

CHAPTER 5. RESULTS ON THE LEARNER FIT AND IMPACT OF CALL CORRECTIVE
FEEDBACK.................................................................................................................................134
  5.1. Learner Fit of Corpus-based and Coded Feedback.........................................................134
    5.1.1. Questionnaire Data ....................................................................................................135
    5.2.2. Semi-structured Interview Data Concerning Impact ..............................................140
    5.2.3. Triangulation of Quantitative and Qualitative Results ............................................146
  5.2. Impact of Corpus-based and Coded Feedback...............................................................147
    5.2.1. Questionnaire Data on Impact ..................................................................................148
    5.2.2. Semi-structured Interview Data on Impact ..............................................................153
  5.3. Triangulation of Quantitative and Qualitative Results ................................................161
  5.4 Chapter Summary...........................................................................................................163

CHAPTER 6. CONCLUSION........................................................................................................166
  6.1. Summary of Findings Related to Each Research Question ........................................166
    6.1.1. Language Learning Potential.....................................................................................166
    6.1.2. Learner Fit ................................................................................................................169
    6.1.3. Impact ......................................................................................................................170
  6.2. Implications for Research, Educational Practice, Pedagogical Design, and CALL Material Evaluation.................................................................171
    6.2.1. Implications for Research ........................................................................................172
    6.2.2. Implications for the Utility of Theoretical Approaches for the Pedagogical
            Design of Corrective Feedback ..................................................................................174
    6.2.3. Implications for the Use of the Logic Model of the Theory of Action for
            Developing a Logical Framework to Evaluate CALL Materials .....................................175
    6.2.4. Implications for Educational Practice......................................................................176
  6.3. Limitations and Suggestions for Future Research ......................................................177
  6.4. Conclusion.....................................................................................................................180

REFERENCES ................................................................................................................................183

APPENDIX A. ERROR AWARENESS SHEET.........................................................................191

APPENDIX B. PRE- AND POST-TESTS..................................................................................192

APPENDIX C. BACKGROUND SURVEY..............................................................................197

APPENDIX D. QUESTIONNAIRE ..........................................................................................199
LIST OF FIGURES

Figure 2.1 Screenshot of written metalinguistic feedback on the articles “a” and “the” (Bitchener, 2008, p. 110) ................................................................. 10

Figure 2.2 Corpus-based comments on partitive expressions and determiners with uncountable nouns (Mendikoetxea, Bielsa & Rollinson, 2011) ......................... 12

Figure 2.3 The Interactionist Approach framework for second-language acquisition process (Chapelle, 1998, p. 23) ............................................................... 26

Figure 2.4 Basic Logic Model of the Theory of Action. (Patton, 2008, p 341). ......................... 33

Figure 2.5 The logical framework of CALL corrective feedback .......................................... 34

Figure 2.6 Logical framework for the development and evaluation of CALL corrective feedback ........................................................................................................ 38

Figure 3.1 The mixed-methods approach with convergent design applied in this study ................ 43

Figure 3.2 Screenshot of My Feedback website ...................................................................... 50

Figure 3.3 Screenshot of a sample of corpus-based feedback in My Feedback ......................... 56

Figure 3.4 Screenshot of a sample of coded feedback ................................................................ 57

Figure 3.5 Screenshot of error correction exercise 1 in My Feedback ...................................... 58

Figure 3.6 Three items about the comprehensibility of the feedback in the questionnaire ........... 61

Figure 3.7 The procedure of this study ..................................................................................... 63

Figure 4.1 Boxplot of error noticing scores on error correction exercise one (Total Scores= 21) ........................................................................................................... 87

Figure 4.2 Boxplot of error noticing scores on error correction exercise two (Total Scores= 21) ........................................................................................................... 88

Figure 4.3 Number of students who perceived the target feedback as helpful for noticing each error type ......................................................................................... 94

Figure 4.4 Boxplot of error correcting scores on error correction exercise one (Total Scores= 21) ...................................................................................................... 101
Figure 4.5  Boxplot of error correcting scores on error correction exercise two ...............102

Figure 4.6  Mean scores for correcting all three types of syntactic errors in exercises one & two .........................................................................................................................103

Figure 4.7  Number of students perceiving the target feedback as helpful for correcting each error type ............................................................................................................109

Figure 4.8  Boxplot of Scores on Pre- and Post-tests for Both Groups .................................................117

Figure 4.9  Mean Scores on Pre- and Post-tests for Both Groups .........................................................117

Figure 4.10  Mean Scores for Correcting All Three Types of Syntactic Errors on Pre- and Post-tests ..................................................................................................................120

Figure 4.11  The most and least difficult errors for correction in the post-test ....................129

Figure 5.1  Boxplot of total scores of each construct concerning the learner fit of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21) ........................................................................................................138

Figure 5.2  Boxplot of total scores of each construct concerning the impact of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21) ......................................................................................151
LIST OF TABLES

Table 2.1 Three cognitive stages in Skill-learning Theory ......................................................... 28
Table 2.2 Approaches to SLA, the attendant design principles of corrective feedback, and the features of corpus-based feedback ................................................................. 32
Table 2.3 Criteria for CALL task appropriateness (Chapelle, 2001 p. 55) ............................. 36
Table 2.4 Research questions and evaluation criteria in this study ........................................ 39
Table 3.1. Demographic information about participants .............................................................. 45
Table 3.2 Sampling design of selected participants in think-aloud protocols and semi-structured interviews ........................................................................................................ 46
Table 3.3 Definitions and main features of the three syntactic errors in this study .................. 49
Table 3.4 Descriptions of websites and programs providing corrective feedback .................. 51
Table 3.4 Continued .................................................................................................................... 52
Table 3.5 Corrective feedback programs/apps and their features .............................................. 53
Table 3.5 Continued .................................................................................................................... 54
Table 4.1 Descriptive statistics of participants’ total noticing scores on exercises one & two (N_{corpus} = 45; N_{code} = 45) ........................................................................ 86
Table 4.2 Descriptive statistics of participants’ scores for noticing three types of syntactic errors in exercises one & two (N_{corpus} = 45; N_{code} = 45) ......................... 89
Table 4.3 Idea units reflecting positive and negative evaluations of both types of feedback on error noticing from both groups (N_{corpus-interviewee} = 9; N_{code-interviewee} = 9) .... 90
Table 4.4 Descriptive statistics of participants’ total error correction scores on exercises one & two (N_{corpus} = 45; N_{code} = 45) ................................................................. 100
Table 4.5 Descriptive statistics of participants’ scores correcting all three types of syntactic errors in exercises one & two (N_{corpus} = 45; N_{code} = 45) ........................................ 103
Table 4.6 Idea Units reflecting positive and negative evaluations of both types of feedback of error correction from both groups (N_{corpus-group-interviewee} = 9; N_{code-group-interviewee} = 9) .................................................................................. 104
Table 4.7 Descriptive statistics of pre- and post-tests (N\textsubscript{corpus} = 45; N\textsubscript{code} = 45) ..................................116

Table 4.8 Descriptive statistics of participants’ scores for correcting all three types of syntactic errors and participants’ gain scores on pre- and post-tests (N\textsubscript{corpus} = 45; N\textsubscript{code} = 45) ..............................................................................................................119

Table 4.9 Reasons for successful error correction in the post-test ..................................................................................121

Table 4.10 Reasons for unsuccessful error correction in the post-test .................................................................123

Table 5.1 Questionnaire items on learner fit and participants’ responses to the Likert-Scale items in the questionnaire ...........................................................................................................................................136

Table 5.2 Total score and standard deviation of each sub-construct on learner fit ....................................................137

Table 5.3 Positive and mixed/negative idea units identified with each theme and subtheme regarding learner fit, as reported in semi-structured interview transcripts ......140

Table 5.4 Questionnaire items on Learner Fit and participants’ responses to the Likert-scale items in the questionnaire ........................................................................................................................................................................148

Table 5.5 Total score and standard deviation of each sub-construct on impact ........................................................150

Table 5.6 Positive and mixed/negative idea units identified with each theme and subtheme regarding impact, as reported in semi-structured interview transcripts ..........153
## NOMENCLATURE

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESL</td>
<td>English-as-a-Second-Language</td>
</tr>
<tr>
<td>L2</td>
<td>Second Language</td>
</tr>
<tr>
<td>CALL</td>
<td>Computer-Assisted Language Learning</td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
</tr>
<tr>
<td>FLT</td>
<td>Foreign Language Teaching</td>
</tr>
<tr>
<td>ZPD</td>
<td>Zone of Proximal Development</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

Developing and evaluating a corpus-based feedback system has been a challenging and rewarding experience for me. This work would not have been possible without the help and guidance of many people. First, I would like to express my earnest gratitude to my major professor, Dr. Carol Chapelle, for her outstanding mentorship, patience, and support during the dissertation process. Thank you for spending hours reading drafts of my dissertation, offering insightful feedback, and helping my intellectual growth throughout all stages of this work. Next, I am deeply indebted to my committee members, Dr. Gulbahar Beckett, Dr. Gary Ockey, Evgeny Chukharev-Hudilainen, and Connie Hargrave. I would like to express my sincerest gratitude to Dr. Gulbahar Beckett, who asked critical questions about my dissertation and offered invaluable suggestions for the refinement of my work. I owe many thanks to Dr. Gary Ockey, who not only reviewed my research design and statistical tests, but also offered significant insight whenever I encountered any difficulties in my dissertation; his feedback strengthened the quality of my dissertation and fostered my knowledge and interest in language testing. Finally, I offer my appreciation of Dr. Evgeny Chukharev-Hudilainen and Connie Hargrave, both of whom were incredibly encouraging and supportive throughout the dissertation.

Additionally, I am deeply thankful to Sarah Davis for helping me revise the corpus-based feedback and select suitable materials for the error correction practice exercises. Without her support, I would have been unable to fulfill my dream of developing corpus-based feedback for use in real language classrooms. I also want to extend my sincerest thanks to my colleagues and friends Yongkook Won, Chuan Jiang, Hong Ma, Roz Hirc, and Amanda Paxton for providing inspiration, assisting in the development of the My Feedback
website, recording oral feedback, and proofreading. The friendships that have arisen in the process of the dissertation are invaluable to me.

Finally, I want to express immense appreciation for my family. My parents, Jun Liu and Xiaobao Chen, continue to support me with their unconditional love. I have shared my successes and failures with them, and their endless support has been the greatest source of positivity and confidence for me throughout my PhD journey. I also express deep thanks to my boyfriend Xiangchen Che, who encourages me in all my pursuits and continuously inspires me to follow my dreams.
ABSTRACT

This dissertation outlines the development of a system of corpus-based feedback and evaluates its effects on students’ error correction and language learning. The design of the corpus-based feedback used in the study was informed by two SLA theories: the Interactionist Approach and Skill-learning Theory. The dissertation theorizes all aspects of the feedback design and discusses the various features of corpus-based feedback from SLA theoretical perspectives. To evaluate the quality of corpus-based feedback, the study compared it to that of traditional coded feedback. Based on the logical framework for evaluating CALL corrective feedback and Chapelle’s evaluation criteria (2001), five research questions were developed concerning three qualities of CALL corrective feedback: language learning potential, learner fit, and impact.

To address the five research questions concerning language learning potential, learner fit, and impact, the study employed a mixed-methods approach to collect quantitative and qualitative data from 90 participants to evaluate and compare the effects of corpus-based and coded feedback. The quantitative data consisted of scores measuring the immediate effects of feedback in error correction exercises, scores measuring learning gains demonstrated by the pre- and post-tests, and students’ ratings of the effects of feedback in Likert-scale questionnaires. The qualitative data consisted of think-aloud exercises and semi-structured interviews, which played an important role in describing participants’ learning experiences with the intervention.

Compared to students using traditional coded feedback, students treated with corpus-based feedback had greater success correcting errors in two error correction practice exercises and transferring knowledge to correct errors in a new text during the
post-test. The benefits of the corpus-based feedback were also confirmed by students in the think-aloud activities and interviews. In addition, the data yielded strong evidence that advanced-low ESL learners found corpus-based feedback to be more appropriate than coded feedback for enabling correction of syntactic errors; students perceived that corpus-based feedback exerted a more positive impact on their affective, cognitive, and intrinsic development towards error correction than the coded feedback. The research methodology and findings of this dissertation make extensive contributions to corrective feedback research and the utility of theoretical approaches in tool development and learning evaluation.
CHAPTER 1. INTRODUCTION

Promoting grammatical accuracy is one of the most important goals in English-as-a-second-language (ESL) academic writing courses. Compared to native English-speaking students, ESL students encounter greater challenges in expressing themselves accurately in their English writing (Hinkel, 2002). One attempt on the part of ESL teachers to address this concern has been to provide second language (L2) learners with corrective feedback in order to promote learning (Bitchener & Ferris, 2012).

According to Lightbown and Spada (1999), corrective feedback is defined as “any indication to the learner that his or her use of the target language is incorrect” (p. 172). In today’s language classrooms, both teacher-written corrective feedback and computer-assisted language learning (CALL) corrective feedback are widely used to correct students’ grammatical errors and improve their grammatical knowledge (Bitchener & Ferris, 2012; Lavolette, Polio & Kahng, 2015). When instructors provide written corrective feedback, they may insert error codes or explanations of errors by hand or by using the track changes feature available through software such as Microsoft Word or other software developed for this purpose. In attempts to improve feedback quality and reduce instructors’ workloads, designers of CALL systems have developed feedback modules that employ web-based technology to provide substantial and immediate corrective feedback.

1.1. Statement of Problems

Teacher-written corrective feedback has long been used in ESL and EFL settings, but the decades-long debates concerning the effects of written corrective feedback on students’ writing continue. On the one hand, numerous studies (Fathman & Wholley, 1990;
Bitchener, 2008; Bitchener & Knoch, 2010; Bitchener, Young & Cameron, 2005; Sheen 2007; Van Beuningen, De Jong & Kuiken, 2008) have provided evidence for the effectiveness of written corrective feedback in promoting student error correction and learning. On the other hand, some researchers have taken a strong stance against employing written corrective feedback in L2 writing courses. For example, Straub (1997) and Truscott (1996) stated that teachers’ inconsistent feedback on grammatical errors, coupled with students’ limited meta-linguistic knowledge, invariably makes corrective feedback futile in language teaching and learning. Moreover, whereas Truscott (2007) argued that teacher-written corrective feedback can be helpful for treating word-level errors, there is little empirical evidence of its help with sentence-level errors.

Unprecedented technological advances are transforming corrective feedback through the introduction and growth of CALL corrective feedback. Many CALL feedback systems have been developed, and many of them provide students with explicit CALL metalinguistic feedback (Cotos, 2011). The debates and challenges related to CALL metalinguistic feedback, however, resonate with more general arguments surrounding written corrective feedback. For instance, the positive effects of CALL metalinguistic feedback on error correction have been demonstrated (Heift, 2004; Heift & Rimrott, 2008; Li, Link & Hegelheimer, 2015), but empirical evidence for its impact on students’ knowledge development, especially syntactic knowledge, is lacking (Warschauer & Ware, 2006).

Research from Second Language Acquisition (SLA) studies in corpus linguistics has further promoted the emergence of new types of CALL corrective feedback and tools. According to McEnery, Xiao, and Tono (2006), a corpus is “a collection of machine-readable
authentic texts which is sampled to be representative of a particular language or language variety” (p. 5). Researchers and pedagogical designers have used insights derived from error analysis with learner corpora to identify persistent errors in L2 learners’ writing and to provide substantial information to language learners (Milton, 2006; Milton & Cheng, 2010; Chuang & Nesi, 2006). Although the benefits of using corpus-based materials and tools for improving grammatical knowledge in L2 writing have been recognized, corpus-based feedback and materials are scarce due to the large amount of work involved in developing such feedback (Granger, 2015). Evaluations of their efficacy in error correction and language learning are also sparse.

1.2. Goals of the Dissertation

Given the gaps in the literature and the challenges of providing high-quality corrective feedback in real ESL classrooms, the preliminary goal of this dissertation was to develop and evaluate a web-based corrective feedback system called My Feedback, which was informed by learner-corpus analysis and was intended to provide systematic, concise, accessible, and well-prepared corrective feedback for students’ error correction. This learning apparatus allowed language learners to refer to specific examples adapted from learner-corpus data, review written and oral explanations of their errors, and receive guidance on making appropriate corrections. Given that the feedback was based on corpus findings and included corpus-adapted examples, it is referred to as corpus-based feedback in this study.

To evaluate the effectiveness of corpus-based feedback on error correction and knowledge learning, a framework based on the Logic Model of the Theory of Action (Patton,
2008) was developed and integrated with criteria for CALL task appropriateness from Chapelle (2001). This study compared the efficacy of the corpus-based feedback from the *My Feedback* system to that of coded feedback, a type of corrective feedback using error codes to help students identify and learn about errors. Employing a mixed-methods research design, this study used triangulated quantitative and qualitative data to evaluate corpus-based feedback and coded feedback.

### 1.3. Significance of the Study

This exploration of the effectiveness of corpus-based feedback in language instruction makes important contributions to SLA research, pedagogical design and evaluation of CALL materials, and pedagogical practices. First, previous corrective feedback research has mainly focused on students’ capacity to revise word-level errors (Bitchener & Ferris, 2012). This study broadens the scope of the current research by analyzing the effects of corpus-based and coded feedback on sentence-level errors. Secondly, the study indicates the significance of collecting both quantitative and qualitative data for evaluating the effects of corrective feedback. Previous research primarily collected quantitative evidence to evaluate students’ performance with the target feedback. This study collected not only quantitative evidence, but also qualitative data to evaluate students’ perceptions of corpus-based and coded feedback. Thirdly, this study exemplifies how to use SLA theoretical approaches to guide the design of corrective feedback and how to develop frameworks for CALL material evaluation. Finally, the project demonstrates great potential in corpus-based feedback for teaching grammatical correction in academic writing courses and reducing teachers’ workloads in such courses.
1.4. Outline of the Dissertation

This dissertation consists of six chapters. This first chapter provided background information about the use of corrective feedback in ESL writing classrooms and discussed a number of unresolved issues concerning the effects of corrective feedback on students’ error correction and knowledge learning. The background information contextualized the research and indicated its significance to the field. Chapter 2 reviews existing literature on the effects of corrective feedback and corpus-based applications for error correction in L2 writing classrooms to identify research gaps. In addition, the chapter discusses the theoretical foundations for the development of corpus-based feedback and the theoretical framework for the evaluation of CALL corrective feedback. Chapter 3 articulates the study's research approach and delineates the participants, teaching contexts, and error types selected, as well as the materials, procedures, and data analysis for each research question. Chapters 4 and 5 present the research findings concerning the learning effects, appropriateness, and future impact of corpus-based and coded feedback. The last chapter synthesizes and evaluates the results and discusses the research implications, limitations, and directions for future studies.
CHAPTER 2. LITERATURE REVIEW

The dissertation describes the development and evaluation of a corpus-based corrective feedback system. This chapter focuses on four key aspects that are fundamental to this investigation. First, it reviews the research on corrective feedback and discusses two central, decades-long controversies surrounding the effects of corrective feedback. The research gaps uncovered in this review leave ample room for further investigation of the effectiveness of corpus-based corrective feedback on syntactic error corrections and language learning. The review goes on to examine studies of the development and evaluation of corpus-based applications and discusses the pedagogical needs and research potential for corpus-based corrective feedback in language classrooms. This section also provides an account of two theoretical approaches to SLA and considers the extent to which they have informed the development of corpus-based corrective feedback. Finally, this section describes the Logic Model of the Theory of Action (Patton, 2008) and Chapelle’s (2001) CALL task evaluation criteria and develops a logical framework for the evaluation of CALL corrective feedback and the formulation of five research questions.

2.1. Corrective Feedback and Issues in Corrective Feedback Research

Corrective feedback has long been used in ESL classroom settings, but the debates surrounding the effects of corrective feedback on error correction and language learning continue unabated (Bitchener & Ferris, 2012). Many studies provide empirical evidence for the effectiveness of corrective feedback in improving revision and writing accuracy in second language learning (Bitchener, 2008; Bitchener & Knoch, 2010; Bitchener, Young & Cameron, 2005; Sheen 2007; Van Beuningen et al., 2008). SLA theories also support the application of corrective feedback for improving students’ writing accuracy. As Schmidt
(1990) suggested, awareness plays a significant role in second-language learning. Given that corrective feedback can contribute to students’ ability to notice their errors by exercising a relatively lower level of awareness and to understand their errors by exercising a relatively higher level of awareness, it should be an effective tool to help students revise their errors and to facilitate language learning (Bitchener & Ferris, 2012). However, some issues concerning the effects of corrective feedback remain controversial. In the following section, I discuss two issues that require further exploration: the effects of different types of corrective feedback on language learning, and the effects of corrective feedback on learning about syntactic errors.

2.1.1. Effects of Different Types of Corrective Feedback on Language Learning

Since the 1980s, an extensive body of research has arisen concerning the relative effectiveness of two types of corrective feedback: indirect and direct (Bitchener & Ferris, 2012; Bitchener & Knoch, 2008; Ferris, 2006; Ferris, 2011; Ferris, Liu, Sinha & Senna, 2013; Ferris & Robert, 2001; Robb, Ross & Shortreed, 1986; Van Beuningen et al., 2008, 2012). Indirect corrective feedback refers to feedback that identifies the existence of an error but does not provide a solution; learners must rely on prior knowledge to make the necessary revision (Bitchener & Ferris, 2012). When instructors provide indirect corrective feedback, they either underline errors or mark them with error codes (e.g. using frag for sentence fragments or ro for run-on sentences). Direct corrective feedback refers to feedback that provides “a correction that identifies where an error has occurred and provides a specific solution to the problem” (Bitchener & Ferris, 2012, p. 131). In many earlier studies, direct feedback was often used interchangeably with the term “direct error correction” (Bitchener & Ferris, 2012). In this type of feedback, instructors cross out errors and provide correct
answers. Metalinguistic feedback, which is primarily used in oral feedback and also used in written feedback to explain the nature of errors and the ways to correct them, is categorized with direct corrective feedback in some recent studies (Lyster & Ranta, 1997; Sheen, 2007).

The conclusions reached by different studies as to the effectiveness of direct and indirect feedback are conflicting. Some studies compare indirect corrective feedback to direct error correction. Researchers (Ferris, 2006; Ferris & Robert, 2001; Lalande, 1982; Robb et al., 1986; Semke, 1984) have reported that indirect corrective feedback provided by highlighting errors or indicating error categories with error codes is as effective as direct error correction in students’ short-term revisions, and showed advantages of indirect feedback over direct error correction on long-term language development.

Accordingly, teachers are primarily advised to provide indirect feedback to students. As those researchers claimed, indirect feedback forces students to be more reflective and analytical about their errors than if they simply transcribed teacher corrections (direct error correction) into the next draft of their papers. Since students are required by indirect feedback to take more responsibility for their scores, they are likely to learn more from the process, to acquire the troublesome structures and to make long-term progress in finding, correcting and eventually avoiding errors. (Ferris, 2011, p. 94)

However, other researchers (Guenette, 2007; Van Beuningen et al., 2012) determined that the findings of some early studies supporting the use of indirect corrective feedback were flawed in their research designs. For instance, in Lalande’s (1982) study, students in the indirect groups participated in more form-focused activities than those in the control group; in Semke’s (1984) research, students in the indirect feedback and direct error correction groups were involved into two different error correction tasks, and the results were not comparable.
Some recent studies (Van Beuningen et al., 2008, 2012) have questioned the effectiveness of indirect feedback on language learning and encouraged teachers to provide direct error correction feedback to students. For example, Van Beuningen et al. (2008, 2012) found that direct error correction and indirect coded feedback were equally effective as editing tools to help students improve the accuracy of an initial text during revision. However, scores in post-tests showed that direct error correction feedback groups significantly outperformed indirect coded feedback and control groups, indicating that direct error correction feedback led to more improved accuracy in new texts. Since the research design in those studies minimized confounding factors (e.g. incomparable execution conditions), their findings are “more robust and compelling” (Bitchener & Ferris, 2012, p. 66).

In addition to direct error correction, there is growing interest in another type of direct corrective feedback: metalinguistic. Sheen (2007) noted that, when teachers provide metalinguistic feedback, they provide metalinguistic explanations, with or without examples, to explain correct forms of words or structures. Based on this definition, Bitchener (2008) and Bitchener and Knoch (2008) developed metalinguistic feedback for their studies. An example of the application of metalinguistic feedback on article errors in Bitchener (2008) is provided in Figure 2.1.

The effects of written metalinguistic feedback on error correction and knowledge development have been compared to those of other types of corrective feedback. When evaluating the relative effects of metalinguistic feedback and direct error corrections, researchers (Bitchener & Knoch, 2008, 2009; Farrokhi, 2012; Van Beuningen et al., 2012;
Sheen, 2007) found that metalinguistic feedback (including examples, explanations and correction forms) is either more effective or as effective in promoting learning as direct error correction. The study (Bitchener & Knoch, 2010) has revealed the superior longitudinal effects of metalinguistic feedback for linguistic development over indirect corrective feedback.

Figure 2.1 Screenshot of written metalinguistic feedback on the articles “a” and “the” (Bitchener, 2008, p. 110).

The strengths of written metalinguistic feedback on L2 writing in the experimental settings have been discussed in many studies (Bitchener & Knoch, 2010; Bitchener & Knoch, 2008, 2009; Farrokhi, 2012; Ferris, 2012; Van Beuningen et al., 2012; Sheen, 2007). However, the use of metalinguistic feedback in real classrooms has been challenged by researchers. For example, Ferris (2010) said that it is unrealistic to expect writing instructors to offer such detailed and high-quality feedback due to large class sizes and time constrictions. Moreover, very few instructors provide corrective feedback with sufficient empirical support; they give feedback mainly based on their intuition and their training experience (Brown, 2012; Ferris, 2011; Hyland & Hyland, 2006).

With technological development, many CALL programs and automated writing evaluation (AWE) systems have been developed to diagnose students’ grammatical mistakes automatically and provide metalinguistic feedback with varying degrees of
explicitness in real classrooms. Some studies examining such systems have mainly discussed the development of the error recognition modules and their feedback systems (Liou, 1991; Reuer, 2003; Tokuda & Chen, 2004). A smaller number involved empirical research exploring the effects of computer-based metalinguistic feedback on paper revision and subsequent writing (Heift, 2004; Heift & Rimrott, 2008; Nagata, 1995). For example, Heift (2004) and Heift and Rimrott (2008) evaluated students’ immediate responses to three types of CALL feedback in an E-tutor system: metalinguistic only, metalinguistic + highlighting, and repetition + highlighting. Results revealed that two types of metalinguistic feedback helped students more than repetition feedback when correcting errors in revision exercises. Another study conducted by Nagata (1995) compared the long-term effects of computer-based metalinguistic feedback with traditional direct error correction. It found that groups who received metalinguistic feedback showed significantly better performance when producing Japanese particles and sentences than groups who received standard error correction in the post-test and delayed post-test. Although explorations of the effectiveness of computer-based metalinguistic feedback are sparse, existing studies have pointed to its positive impact on students’ error correction and knowledge learning (Heift, 2004; Heift & Rimrott, 2008; Nagata, 1995).

An emerging trend in metalinguistic feedback involves the incorporation of learner corpora into the feedback; such feedback has been termed “corpus-based” (Seliem & Ahmed, 2009). Corpus-based feedback integrates sentences from corpora and research findings from corpus analysis to facilitate error understanding and language learning. In the sample provided in Figure 2.2, corpus-based feedback includes examples from learner
corpora with corresponding corrections and explanations to illustrate the use of partitive expressions and determiners with uncountable nouns.

![Partitive expressions and determiners with uncountable nouns](image)

**Figure 2.2 Corpus-based comments on partitive expressions and determiners with uncountable nouns (Mendikoetxea, Bielsa & Rollinson, 2011).**

The benefits of corpus-based materials for teaching L2 grammatical knowledge have been discussed in some studies, but empirical research on the implementation of corpus-based corrective feedback is scarce (Granger, 2015). Therefore, the dissertation was predicated on developing a corpus-based feedback system and evaluating its effectiveness for teaching error detection and correction in L2 writing. The evaluation of corpus-based corrective feedback provided both quantitative and qualitative evidence concerning the use of this new type of metalinguistic feedback in real ESL classrooms.

### 2.1.2. Effects of Corrective Feedback on Correcting Syntactic Errors

In addition to the concerns surrounding the efficacy of different types of corrective feedback, another criticism of corrective feedback involves its efficacy for teaching students to avoid and correct syntactic errors. Truscott (2007) has claimed that corrective feedback can only be helpful treating errors that “are relatively simple and can be treated as discrete items” (2007, p. 258). As for syntactic structures, which are always complex and comprise a
series of discrete items, written corrective feedback has been of little help (Truscott, 2007). Many studies have evaluated either the immediate or long-term effects of corrective feedback on specific discrete items (Bitchener & Knoch, 2008, 2009; Chuang & Nesi, 2006; Farrokhi, 2012; Ghandi and Maghsoudi, 2014; Sheen, 2007; Van Beuningen et al., 2012). For example, Ghandi and Maghsoudi (2014) explored the effect of direct and indirect feedback on Iranian EFL high school learners’ spelling errors. Sheen (2007) investigated the effects of different direct types of corrective feedback on the use of indefinite and definite articles. The findings of these studies suggest the effectiveness of specific types of corrective feedback on developing knowledge about target error categories.

A small number of studies have discussed the efficacy of corrective feedback on the correction of syntactic errors, along with other types of errors, in students’ drafts (Ferris, 2006; Ferris & Robert, 2001; Sachs & Polio, 2007; Van Beuningen et al., 2008, 2012; Li and Hegelheimer, 2013). For example, Ferris and Robert (2001) explored the effects of different types of indirect corrective feedback on different error categories: verb errors, noun-ending errors, article errors, word-choice errors, and syntactic errors (i.e. sentence fragments, run-ons, and sentence structure errors). They found both types of indirect corrective feedback (codes and highlighting) to be less helpful for correcting syntactic errors than they were for errors in the four rule-governed, discrete error categories. The study did not explore the reasons behind its finding.

In another article discussing the impact of feedback on self-editing of word- and sentence-level errors in L2 writing, Li and Hegelheimer (2013) discussed the CALL metalinguistic feedback provided by the mobile-assisted grammar application Grammar Clinic. Students were first required to practice error identification and correction with
metalinguistic feedback. After the treatment in *Grammar Clinic*, students used acquired knowledge to correct five types of errors: article errors, preposition errors, verb errors, run-on sentences, and sentence fragments. The analysis of error rates between the first and final assignments showed a sharp decrease in the rates of run-on sentences, verb errors and preposition errors between the two assignments, yet there was no statistically significant difference in the rate of sentence fragments and article errors in the two papers. Researchers noted that the reduction of run-on sentences in students' papers “is very encouraging given the difficulty of this sentence-level error” (p. 145). However, the paper did not venture to explain the phenomenon.

According to Ferris, Chaney, Roberts, and McKee (2000, cited in Bitchener & Ferris, 2002), sentence fragments, run-on sentences, and sentence structure errors are among the 15 most common errors in ESL writing (ranked 13th, 10th, and 1st, respectively). In contrast to the extensive research into sentence-level errors, the research on the effects of corrective feedback on these syntactic errors is very limited. To my knowledge, there has been no study exclusively focusing on sentence-level errors. Given the frequency of syntactic errors in students’ papers and the limited attention to those errors in the literature, the dissertation is going to assess the effects of corrective feedback on the correction of three syntactic errors: sentence fragments, run-on sentences, and subordinate clause errors. Quantitative and qualitative data were triangulated to provide information on the effects of corpus-based and coded corrective feedback on students’ error corrections and knowledge development of syntactic errors.

According to the discussion in the first part of the literature review, the unsettled debate surrounding the effectiveness of corrective feedback indicated the importance of
this topic in second language research. The dissertation focused on the potential for corpus-based corrective feedback to aid students in correcting syntactic errors and learning syntactic knowledge. In the next part, studies on corpus-based applications, including corpus-based corrective feedback, are reviewed and their effects on error correction and language learning are investigated.

2.2. Learner Corpora and Corpus-based Tools for Error Correction in L2 Writing Classrooms

Learner corpora comprise “collections of authentic foreign language/second language textual data assembled according to explicit design criteria for a particular SLA/FLT purpose” (Granger, 2002, p. 7). To date, more than one hundred written and spoken English-language-learner corpora have been developed (Cotos, 2014). Along with this trend, an increasing number of studies with learner corpora have been conducted to gain a better understanding of L2 learners and their second language acquisition processes. For example, analyses with large- or small-scale learner corpora have revealed typical errors that learners have had in language learning. In addition, some researchers have conducted comparative analyses of the different uses of lexical and syntactic structures by L2 learners at different proficiency levels, or of the over- or under-use of target linguistic features between native and non-native writers. Analyses of various learner corpora have provided insight into the stages of learning a new language and the specific difficulties attendant on those stages. Findings from learner corpus research are therefore invaluable. They shed light on the language control of learners at different acquisition stages and indicate the importance of developing learning materials to address students’ real needs in language learning.
In L2 writing, the impact of learner corpora is noticeable in the development of learner dictionaries and ESL textbooks. Since the 1990s, various learner dictionaries, such as the Longman Language Activator (1993), the Macmillan English Dictionary for Advanced Learners (2002), and the Cambridge Advanced Learners' Dictionary (2013) have been developed. These dictionaries not only include definitions of each word but also address errors that occur frequently in the work of learners from a variety of L1 backgrounds. Readers learn those errors through reviewing authentic corpus-based examples and are given suggestions for error correction. The development of L2 writing courses is another example of learner corpus applications. Informed by research findings, textbooks such as Writing Clearly (2011) and the Grammar and Beyond series (2012) emphasize specific errors in ESL writing in each unit. They provide error warnings for each type of error, explicitly teach grammatical knowledge, and enable students to apply knowledge through writing tasks.

In the technological era, the exploration of the use of learner corpus data and findings to teach grammatical errors in L2 writing has been further expanded with corpus-based applications. Such computer-assisted language tools have integrated corpus resources and web-based technologies to facilitate grammatical knowledge learning. Given that the main goal of the dissertation was to develop and evaluate a CALL corpus-based corrective feedback system, previous studies of the development of CALL corpus-based tools were reviewed. The evaluation of the effectiveness of these tools in classroom settings were discussed and the research gaps were identified.
2.2.1. Direct and Indirect Use of Learner Corpus Data in CALL Tool Development in L2 Writing Classrooms

According to Maingay and Rundell (1987), learner corpus analyses that reveal learners’ difficulties can help anticipate and identify grammar errors in learners’ writing. To meet the needs of students at specific language proficiency stages or with special L1 backgrounds, many instructors and researchers undertake the role of pedagogical designers to develop CALL writing tools with learner corpora (Mendikoetxea et al., 2010). According to Meunier (2010), learner corpus data have been used either directly or indirectly to address typical grammar errors in L2 writing classrooms. The CALL tools that make direct use of corpus data use a collection of students’ essays from learner corpora, without any modification, for teaching and practice to promote L2 writing. Conversely, the ones that make indirect use of corpus data employ findings from corpus analysis to inform the development of teaching materials or select target erroneous sentences as examples for instruction (Meunier, 2010).

The direct use of learner texts in L2 writing courses is very rare due to concerns that students may be overexposed to various L2 errors. One study conducted by Hegelheimer and Fisher (2006) described a corpus-based, database-driven application, iWrite, used in ESL writing classes at a U.S. university. To develop this system, 45 essays from an English placement test (EPT), comprising 12,839 words, were analyzed, and 1268 grammatical and lexical errors were marked. In the iWrite web-based system, students accessed all the marked errors in all the essays and referred to the appropriate corrections and explanations. They could also select one type of error, such as misspelling, and read all the instances of that error across the learner corpus. The learner corpus data were also used to
generate exercises: when an error category was selected, the appropriate errors across all the texts were highlighted and Microsoft Word worksheets for error correction exercises were created for downloading. The iWrite system increased learners’ error awareness, taught various types of errors in context, and promoted the development of error correction exercises for in-class activities.

The L2 writing classroom most commonly makes indirect use of learner corpora (Granger, 2015). Instead of directly using students’ essays as teaching materials, CALL tools are informed by the findings of corpus research to develop lessons and exercises (Hegelheimer, 2006; Milton, 2006; Milton & Cheng, 2010; Chuang & Nesi, 2006; Cowan, Choo, and Sunny, 2014). For example, GrammarTalk (Chuang & Nesi, 2006) is an online program based on the error analysis of a corpus of 50 assignments composed by Chinese pre-undergraduates in a British university. The systematic corpus-based analysis (Chuang & Nesi, 2006) has revealed that the mismanagements of English articles and prepositions were the most recurrent grammatical errors in Chinese students’ writing assignments. Accordingly, GrammarTalk has developed two units on those error categories. In each unit, students learn different aspects of the target linguistic feature in 7-8 tutorials and then test their knowledge in grammar consciousness-raising and production-oriented tasks. In those tasks, GrammarTalk provides answers and explanations to reinforce relevant grammatical rules and to help students become aware of the gap between their interlanguage and the standard forms of the target language.

Learner corpus data have also been used indirectly to inform the development of corrective feedback for students’ writing and to optimize feedback design in CALL programs. Mark My Words (Milton, 2006) and Check my Words (Milton & Cheng, 2010) are
two corpus-based feedback systems. The developers of both systems analyzed English texts from a large learner corpus, extracting more than 500 typical lexico-grammatical and style errors and composing feedback for each type of error. The feedback included a brief metalinguistic explanation of the target error and hyperlinks to relevant language sources with correct forms and multimedia materials. When grading papers, rather than writing comments for each individual error, instructors can draw from a library of pre-written feedback and insert feedback into student papers accordingly. Using the corrective feedback, learners are expected to become sensitized to common lexico-grammatical errors and learn how to correct them.

To date, a growing number of studies discuss CALL applications that have been developed with direct and indirect use of learner corpus texts (Hegelheimer, 2006; Milton, 2006; Milton & Cheng, 2010; Chuang & Nesi, 2006; Cowan, Choo, and Sunny, 2014). Built on solid research findings, the teaching materials, exercises, and corrective feedback in those applications focus on recurrent errors in ESL students’ writing (Granger, 2015). Language learners increase their error awareness through explicit lessons on common errors and error warnings rather than a comprehensive study of grammar. Moreover, technology has enlivened the learning experience. Rich online resources and multimedia learning materials have helped to deliver grammatical knowledge in an engaging and interactive way. In addition, teaching materials and corrective feedback use authentic examples from learner corpora to explain the nature of errors clearly and accessibly.

Although corpus-based tools boast features that promote grammatical knowledge learning, the large amount of error analysis and web development entailed in their design has kept their numbers low. In light of the value of learner corpus research and its CALL
applications, I created a web-based corrective feedback system making indirect use of a local learner corpus. Analyzing 90 essays in an English Placement Test (EPT) corpus, I designed a corpus-based corrective feedback system with the assistance of many colleagues in the ISU English Department and developed this system with a student in the Department of Computer Science. The dissertation focused on the development and evaluation of the system.

### 2.2.2. The Evaluation of Corpus-based Pedagogical Materials and Tools in CALL

As previously mentioned, an increasing number of studies discuss CALL tools that have been developed with direct and indirect use of learner corpus texts (Hegelheimer, 2006; Milton, 2006; Milton & Cheng, 2010; Chuang & Nesi, 2006; Cowan, Choo, and Sunny, 2014). They detail the data analysis process and findings with learner corpus data and tool development procedures. The quantitative and qualitative evidence for the effectiveness of those tools on language learning, however, is insufficient.

Some research has entailed quantitative analysis of learning outcomes from the use of corpus-based tools in real classroom settings (Chuang, 2017; Hegelheimer, 2006; Milton and Cheng 2010), but the conclusions based on the findings were limited due to small sample sizes. Hegelheimer (2006) evaluated the effects of *iWrite*, a corpus-based application, on the improvement of error awareness and writing accuracy for nine undergraduate ESL students. The descriptive statistics of error percentages from paper 1 to paper 5 indicated the effectiveness of *iWrite* in helping students transfer knowledge of syntactic and lexical errors to their own papers. Chuang (2017) also conducted an empirical study to evaluate the effectiveness of a corpus-based application, *GrammarTalk*, for teaching the use of English definite and indefinite articles to Chinese undergraduates.
Scores from pre-, post-, and delayed post-tests showed that the 14 learners in the experimental group using GrammarTalk and the 10 learners in the control group not using the application performed similarly in the pre-test and two post-tests. However, the experimental group did show a significant improvement between the pre- and post-tests, while such improvement was not observed in the control group. In summary, both small-scale studies found positive effects of corpus-based feedback on prompting students to correct grammatical errors.

In another study, Milton and Cheng (2010) described the features of Check my Words, a corpus-based corrective feedback system, and discussed its classroom implementation and assessment. In the article, Milton and Cheng (2010) claimed that the experimental group that received corpus-based feedback from the application showed more significant improvement in their writing accuracy and fluency when revising their own writing than the control group that did not received corpus-based feedback. However, this claim was not supported with any quantitative evidence.

To my knowledge, the only study that has provided an evaluation of a corpus-based tool with an adequate research design and sample size was conducted by Cowan, Choo, and Lee (2014). They evaluated the efficacy of the tool ESL Writing Tutor in improving Korean-speaking leaners’ knowledge of persistent errors in their English writing. The ESL Writing Tutor applies analysis and insights from a large learner corpus produced by Korean undergraduate and graduate students and develops exercises focusing on four types of errors: pseudo-passive errors, article errors, ergative verbs, and noun phrases with modifiers. In each unit, explicit lessons and multiple error correction exercises with metalinguistic explanations or hints were given to students for practice and language
learning. The forty students in the experimental group received exercises and feedback from the system, while the eighteen students in the control group did not. The pre-test/post-test/delayed-post-test scores revealed that the experimental group made significant improvement in all four error categories between the pre-test and the first post-test (administered five weeks after the pre-test) and in three error categories between the pre-test and the delayed post-test (administered five months later). By contrast, students in the control group who received no feedback showed no significant differences among their pre-test, post-test, and delayed post-test scores. The authors advocated integrating selected sentences from learner corpora into exercises and feedback, and they commended the effectiveness of the tool to address persistent errors.

In addition to assessing learner outcomes with corpus-based tools, some research has also investigated the perceptions of students and instructors of the effectiveness of corpus-based applications. Hegelheimer (2006) used questionnaires and semi-structured interviews to assess students’ learning experiences with iWrite. Answers to questionnaires revealed that learners generally believed that error descriptions and in-class error correction activities were beneficial to their writing and grammatical knowledge; they hoped to continue using this program out of class in the future. During the semi-structured interviews, students articulated the reasons that they enjoyed the program. After interacting with iWrite, students believed that they had increased their grammatical awareness of the common errors in their own writing and understood how to improve their own writing. Similarly, when discussing another corpus-based application, Chuang (2017) collected answers from surveys to evaluate Chinese undergraduates’ satisfaction with their learning experience. Students generally believed that the application was useful
because the instructions were clear and the information from the feedback was comprehensible. Also, they felt that the tool facilitated self-exploration and independent learning; the flash and videos in the applications made their learning experience fun and interactive.

Reviewing the literature has revealed improved learning outcomes and overall positive learning experiences with corpus-based tools. At the same time, problems in current studies on the evaluation of corpus-based tools are salient. First, few studies have included the combination of quantitative and qualitative analyses that would achieve a holistic evaluation of those applications. Some research (Cowan et al., 2014) has relied on quantitative analysis to assess the learning outcomes with the tools, while some research (Milton, 2006; Milton & Cheng, 2010) does not include any quantitative or qualitative evidence. Only two studies (Hegelheimer, 2006; Chuang, 2017) explore students’ learning outcomes and users’ perceptions with quantitative and qualitative analysis. Including both forms of analysis provides a well-rounded picture of the effectiveness of tools in the learning environment. Second, all the current studies focus on the learning outcomes with applications; none of them have included learning process data to assess students’ performance and perceptions. According to Chapelle (2003), “learning process data” refers to “records of learners’ language and behavior documented while they are working on computer-mediated tasks” (p. 98). Including learning process data allows researchers to gain insight into how students interact with the CALL applications and what difficulties students encounter when using the tool. Third, a majority of the existing studies have relied on a very small number of participants. For example, Hegelheimer (2006) surveyed nine participants, and Chuang (2017) had eleven.
To fill these research gaps, the dissertation collected quantitative and qualitative evidence to evaluate the effects of a corpus-based corrective feedback system. Quantitative data were collected via error correction practice exercises, pre- and post-tests, and questionnaires, while qualitative data were gathered via think-aloud protocols, and semi-structured interviews. Data from think-aloud protocols were analyzed to illustrate students’ interactions with corpus-based feedback in the learning process. Error correction exercise scores, pre- and post-test scores, semi-structured interview and questionnaire data were used to explore learner output, learning outcomes, and students’ perception of learning. Moreover, this study recruited 90 participants, with 45 in the code group and 45 in the corpus group, to compare the effects of corpus-based and coded feedback. Increasing the sample size allowed for a more precise estimate of the treatment effects compared to previous studies.

In sum, the potential pedagogical needs of students learning about grammatical errors in L2 writing using corpus-based applications have been discussed in previous studies. However, corpus-based tools, including corrective feedback systems, are very limited. Moreover, the quantitative and qualitative evidence to assess these tools is insufficient. Driven by pedagogical and research needs, a corpus-based feedback system has been developed by the primary researcher, with the assistance of her colleagues in English Department and a PhD student in the Computer Science Department, for use in the L2 writing classroom. The next two sections articulate the theoretical foundations for the development of CALL corpus-based corrective feedback and construct a theoretical framework for evaluating CALL corrective feedback.
2.3. Theoretical Foundations for the Development of Corpus-based Corrective Feedback

Over more than two decades, various theoretical approaches to second language learning have been explored. When developing CALL tools and activities, researchers and instructors have combined different approaches to SLA as theoretical foundations of their design. As Chapelle (2009) suggested, when developing and evaluating CALL activities and tools, it is limiting to draw solely on one theoretical approach to SLA because “[e]ach theory focuses on a set of phenomena, whereas CALL activities can span to a broad range of learning opportunities” (p. 747). Combining theoretical approaches takes account of various factors that contribute to successful language learning in CALL contexts (Chapelle, 2009). This dissertation uses two theoretical approaches to SLA—the Interactionist Approach, and Skill Learning Theory—to explain second language learning from psycholinguistic, and human cognitive perspectives and to delineate the key features of high-quality CALL corrective feedback. These theoretical approaches provide solid bases for the development of CALL corpus-based corrective feedback.

2.3.1. The Interactionist Approach and Corrective Feedback

Many researchers in second language learning have articulated theoretical perspectives regarding the role of corrective feedback (Bitchener & Ferris, 2012). One theoretical position that lends great support to error treatment through corrective feedback is the Interactionist Approach. The Interactionist Approach, which is closely interwoven with the Noticing Hypothesis (Schmidt, 1990, 1995, 2001) and the Output Hypothesis (Swain, 1985, 1995, 2005), constitutes an important model in SLA research (Ramirez, 2005). Gass (1997, cited in Chapelle, 1998) outlined a framework for second-
language acquisition that identifies six fundamental stages involved in the conversion of input into output in the Interactionist SLA Approach (see Figure 2.3).

![The Interactionist Approach framework for second-language acquisition process (Chapelle, 1998, p. 23).](image)

The Noticing Hypothesis emphasizes two different types of attention practiced by learners: “noticing” and “understanding,” which Interactionists refer to as “apperception” (stage 2) and “comprehension” (stage 3), respectively. As Schmidt (2001) articulated, SLA “is largely driven by what learners pay attention to and notice in target language input and what they understand the significance of noticed input to be” (p. 4). Even though researchers have recognized the value of acquiring a language through an implicit and subliminal method, they recognize that this method is neither practical nor effective for second language acquisition (Schmidt, 1995; Smith, 1991). The provision of corrective feedback is supported by the Noticing Hypothesis: feedback facilitates learners’ ability to notice the gap between their interlanguage form and the target form and to understand the target grammatical rule, thus promoting interlanguage development. If learners are given feedback, such as bolding or metalinguistic explanations, as input enhancements to heighten the saliency of target L2 linguistic features, learners’ attention to a linguistic problem creates the potential for error correction and effective language learning (Doughty & Williams, 1998; Sheen, 2011). Moreover, when the input is pre-modified through lexical and syntactic simplification and elaboration (Chapelle, 1998; Larsen-Freeman & Long,
1991), comprehensibility is increased. The Noticing Hypothesis not only unequivocally supports the provision of corrective feedback to L2 learners, but also suggests potential methods by which to develop effective corrective feedback. Effective feedback should constitute input enhancement in two ways: 1) the target linguistic feature in the feedback should be made salient in order to direct learners’ attention to the target problems; and 2) spoken and written language in corrective feedback should be simplified or modified to facilitate learners’ ability to understand target linguistic norms and formulate new hypotheses.

The Interactionist Approach also draws heavily on the Output Hypothesis developed by Swain (1985, 1995, 2005) (see Figure 2.3), which further justifies the use of corrective feedback in L2 learning. Swain argued that learners’ comprehension of input cannot guarantee their acquisition of linguistic knowledge; he (1995, 2005) further discussed the importance of “intake” and “integration” in the process of language learning. After receiving apperceived and comprehended input, learners make a new hypothesis concerning linguistic forms, and this hypothesis temporarily becomes part of their language system (Stage 4: Intake) (Chapelle, 1998). Through repeated practice or exposure to exemplars, the new hypothesis is confirmed and integrated into learners’ existing language-knowledge systems (Stage 5: Integration). Swain (1985) found that learners are not always able or self-motivated enough to develop learner intake and knowledge integration. He said that corrective feedback can be used to push output from L2 learners, allowing them to test their new hypothesis in their intake and enabling them to integrate new knowledge and modify their existing language system (Swain, 1995, 2005). Learners are informed and motivated by corrective feedback to stretch their interlanguages to produce accurate
output. As a result, in addition to attention-directing and comprehension enhancing, another significant feature of effective feedback is output-inducing.

2.3.2. Skill-Learning Theory and Corrective Feedback

Another theoretical approach that integrates cognitive psychology into SLA research and informs the role of corrective feedback in L2 learning is Skill-Learning Theory. Originating from the information-processing model developed by McLaughlin (1987), Skill-Learning Theory explains general human learning processes (Chapelle, 2009). McLaughlin believed that learning a second language involves acquiring various cognitive skills, a process that requires the automatization of component sub-skills through practice. Anderson (1993) and Leeman (2007) furthered McLaughlin's (1987) ideas and discussed three stages and processes through which L2 learners acquire second language knowledge, as shown in Table 2.1.

Table 2.1 Three cognitive stages in Skill-learning Theory.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Controlled phase I: Formal instruction</td>
<td>The cognitive stage</td>
<td>Acquisition of declarative knowledge</td>
</tr>
<tr>
<td>Controlled phase II: practice</td>
<td>The associative stage</td>
<td>Proceduralization</td>
</tr>
<tr>
<td>Automatized phase</td>
<td>The autonomous stage</td>
<td>Automatization</td>
</tr>
</tbody>
</table>

During what Anderson calls the cognitive stage, learners acquire declarative knowledge—the explicit grammatical rules of a specific language feature—using relatively high levels of attention and effort to understand and memorize this knowledge. The second, associative stage describes the shift from controlled to automatic processing of knowledge. Here, learners try to enact the rules in their productions; in this process, the declarative
knowledge about those grammatical rules is tested and activated repeatedly, thus being gradually converted into less explicit procedural knowledge that requires less attention to apply. The third stage represents the full mastery of linguistic knowledge. As the term “autonomous stage” indicates, language knowledge has at this stage been integrated into the learner’s language system and is used automatically, requiring minimal conscious attention.

Skill-Learning Theory emphasizes the significance of corrective feedback in language learning and knowledge reinforcement. First, this theory emphasizes the roles of knowledge instruction and practice in helping develop declarative knowledge. This juncture is where corrective feedback comes into play. Leeman (2007) said that corrective feedback provides students with declarative knowledge of a linguistic form or indicates “the need to change the scope of a given rule or procedure” (p. 117). In addition, corrective feedback segments L2 production tasks and skills into smaller ones, making the learning objectives clear and manageable. A writing task requires L2 learners’ attention to various factors, which may lead to cognitive overload, grammatical errors, and logical difficulties. Corrective feedback can help narrow the scope of the revision tasks, prioritize students’ attention to a specific grammatical problem, and eliminate distracting factors (Carroll, 2001; Leeman, 2007).

2.3.3. Implications of SLA Theories for Pedagogical Design of Corrective Feedback

The pedagogical design used in this study is informed by the Interactionist Approach, and Skill-learning Theory. In the Interactionist Approach, “interactive work and negotiation work... connect input, internal learner capacities, particularly selective
attention, and output in productive ways” (Long, 1996, pp. 451-452). Pawlak (2013) explained that interactive negotiation work includes both negotiation of the meaning of linguistic features that interfere with communication and negotiation of a problematic language form in oral communication or written discourse. Negotiation of form relates to “the issues of how competent speakers react to leaners’ language errors” (Gu, 2018, p. 16). When discussing the Interaction Hypothesis, Long (1996) stated that negotiation of form “may be facilitative of L2 development, at least for vocabulary, morphology and language-specific syntax and essential for learning certain specifiable L1-L2 contrast” (p. 414). Many researchers (Lyster, 1998; Pawlak, 2013; Sheen, 2011) perceived corrective feedback as a vehicle for the negotiation of form between instructors and L2 learners. Supported by the Interactionist approach, corrective feedback provided about language should: 1) allow the target linguistic feature(s) in the input to be salient, 2) ensure semantic and syntactic comprehension, and 3) motivate students to produce “pushed-output.”

The Interactionist Approach guided the development of the corpus-based feedback used in this study in three ways. First, color-coded metalinguistic feedback is provided to facilitate students’ attention to errors and understanding of the input. Many research papers extol the potential of metalinguistic feedback to promote learners’ error awareness and understanding with descriptions and explanations (Chapelle, 1998; Gholaminia, Gholaminia & 2014; Shintani & Ellis, 2013). In the My Feedback system, students interact with the feedback, negotiating problematic language forms and developing their error awareness and comprehension. Second, the feedback provided by My Feedback is output-inducing. Elicitation questions in the My Feedback system such as, “Now, can you find the problem in your sentence?” are used in oral feedback to elicit information from students,
thus enhancing two-way communication in the error correction activity and inducing output.

The use of metalinguistic feedback provided by *My Feedback* is theoretically supported not only by the Interactionist Approach, but also by Skill-Learning Theory. Skill-Learning Theory emphasizes the importance of explicit instruction in language learning. Before L2 learners know how to use specific grammatical structures, they acquire knowledge about target linguistic features at the first cognitive stage of learning. At this stage, corrective feedback plays an important role in introducing explicit knowledge and grammatical rules. This is how corpus-based feedback supports students’ development of declarative knowledge of syntactic errors. In a piece of corpus-based corrective feedback, metalanguage—grammatical terms such as “subject” or “main clause”—are highlighted next to a semantically or morphologically related item in the example sentence to explain abstract rules in context. What is more, the feedback includes a grammatical rule section, with a clear explanation of why the authentic learner examples included are problematic.

In conclusion, two approaches to SLA—the Interactionist Approach with the Noticing and Output Hypotheses, and Skill-Learning Theory—constitute the theoretical foundations for the design of the corpus-based feedback in this project, as summarized in Table 2.2. Informed by these principles listed in the table, I have devised the corpus-based feedback to help students 1) notice the gap between the interlanguage form and the correct form and understand errors with colors, 2) understand target linguistic forms with examples and metalinguistic explanation, 3) produce “pushed-output”, and 4) learn and reinforce declarative knowledge with oral and written metalinguistic feedback through
repeated practice. In the next section, the theoretical framework for evaluating CALL corrective feedback is articulated.

Table 2.2 Approaches to SLA, the attendant design principles of corrective feedback, and the features of corpus-based feedback

<table>
<thead>
<tr>
<th>Approaches to SLA</th>
<th>Design Principles of Corrective Feedback</th>
<th>Features of corpus-based feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Interactionist Approach</td>
<td>• allows the target linguistic feature(s) in the input to be salient</td>
<td>• Color-coded feedback</td>
</tr>
<tr>
<td></td>
<td>• facilitate learners’ ability to understand target linguistic norms and formulate new hypotheses</td>
<td>• exemplar-based feedback</td>
</tr>
<tr>
<td></td>
<td>• motivates students to produce “pushed-output.”</td>
<td>metalinguistics feedback</td>
</tr>
<tr>
<td></td>
<td>• introduces explicit knowledge and grammatical rules to develop declarative knowledge;</td>
<td>• Oral feedback inducing output</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• written and oral metalinguistic explanation in the feedback</td>
</tr>
</tbody>
</table>

2.4. Theoretical Framework for the Evaluation of CALL Corrective Feedback

In addition to the development of the CALL corpus-based corrective feedback, another main goal of this study is to evaluate CALL corrective feedback and its use in language classrooms. The Logic Model of the Theory of Action (Patton, 2008) and Chapelle’s (2001) evaluation criteria for CALL task appropriateness are combined to develop the Theory of Action framework for the CALL corrective feedback evaluation in this study. Grounded in SLA theories, this framework connects the underlying theoretical background of the pedagogical design of CALL materials to the evaluation objectives of this project. In the following section, I discuss the Logic Model (Patton, 2008) and Chapelle’s evaluation criteria (2001) and how they are integrated to develop the evaluation framework for CALL corrective feedback. I also articulate how the framework informs the construction of the research questions in this study.
2.4.1. The Logic Model of the Theory of Action and the Development of the Logical Framework for Evaluating CALL Corrective Feedback

The Theory of Action, derived from action research, originated in the context of social theory (Coleman, 1986) and has recently been accepted as an appropriate tool for language program evaluation (Mathison, 2011; Norris, 2016). The Theory of Action originally emphasized the distinction between espoused theory and theory-in-use (Argyris, 1980). Espoused theories are what practitioners believe to be true about their actions; theory-in-use denotes the real consequences of actions, often unforeseen by practitioners. As Patton (2008) notes, “In this conundrum of dissonance between stated belief and actual practice lies a golden opportunity for reality tests” (p. 339). A Theory of Action framework for evaluating language programs and tools describes a process of putting stated belief into an empirical test. A basic Logic Model of the Theory of Action (Patton, 2008), as seen in Figure 2.4 below, constructs a means-ends hierarchy.

The first components in a chain of actions, comprising inputs/resources, constitute materials that are applied in activities. Those activities are actual interventions: during the activities, participants interact with inputs/resources and produce outputs; after the activities, they apply knowledge acquired from the activities to solve other problems, revealing short-term outcomes and long-term impacts.

![Basic Logic Model of the Theory of Action](image)

Figure 2.4 Basic Logic Model of the Theory of Action. (Patton, 2008, p 341).
The first components in a chain of actions, comprising inputs/resources, constitute materials that are applied in activities. Those activities are actual interventions: during the activities, participants interact with inputs/resources and produce outputs; after the activities, they apply knowledge acquired from the activities to solve other problems, revealing short-term outcomes and long-term impacts. The arrows among the five components make up a continuous series of actions and indicate a sequential logical relationship among input, activities, expected outputs, outcomes, and impacts. Norris (2016) suggested that such a logical model could be adapted for evaluating language programs, since the model “outlines the programs’ theory of change by spelling out the critical assumptions that define all identifiable components of a program” (p. 177). Gruba, Cardenas-Claros, Suvorov, and Rick (2016) have already used evaluation argument for CALL evaluation. In this study, a logical framework has been established to identify the key theories for CALL corrective feedback development and the key assumptions for its evaluation (Figure 2.5).

![Logical framework of CALL corrective feedback](image)

Figure 2.5 The logical framework of CALL corrective feedback.

The framework includes six components involved in the various stages of developing and implementing CALL corrective feedback and in its immediate and long-term hypothesized effects and impact on learning. Specifically, theoretical principles (the first component) inform the pedagogical design of CALL corrective feedback (the second component). Then, the CALL corrective feedback is integrated as a part of the activities (the
third component). Through interacting with feedback in the activities, students are expected to produce immediate learning outcomes (the fourth component). The assumed immediate learning effect is anticipated to lead to learning gains in the long term (the fifth component), which will bring about a positive impact on language learning (the sixth component).

Building a logical framework for program and tool evaluation is valuable (Norris, 2016). A logical framework specifies all the actions and hypotheses in a chain that lead to the achievement of a given ultimate impact on language learning. According to Norris (2009), the logical framework approach collects evidence regarding 1) whether all the hypotheses are realized as intended, 2) what factors constrain or foster the realization of each hypothesis within a given learning context, and 3) whether the use of the language tool/program for learning achieves its desired ultimate goal. Using the logical framework, this study is intended to evaluate the effectiveness of CALL corpus-based corrective feedback in language learning; this framework offers a format for connecting hypotheses with evidence. In the next section, the criteria used to evaluate those hypotheses and collect evidence are delineated.

2.4.2. Chapelle’s Criteria for CALL Task Appropriateness and the Development of the Logical Framework for Evaluating CALL Corrective Feedback

Chapelle’s (2001) criteria for CALL task appropriateness are employed to develop evaluation criteria of hypotheses in the logical framework for evaluating CALL corrective feedback. In Computer-assisted Language Learning, Chapelle (2001) proposed six criteria for evaluating the appropriateness of a given CALL task for language learning: language-learning potential, learner fit, meaning focus, authenticity, positive impact, and practicality
(Table 2.3). Informed by SLA cognitive and social-affective approaches and research, these criteria capture different dimensions of language learning when used to evaluate CALL materials. For example, informed by psycholinguistic approaches that emphasize attention in language learning, Chapelle (2001) established language-learning potential as one criterion to evaluate learners’ attention to target word forms with CALL tasks. Moreover, Chapelle (2001) advocated conducting quantitative and qualitative analyses for CALL task evaluation. The quantitative and qualitative data are complementary, providing evidence for evaluation and ultimately contributing to a fuller picture for CALL task evaluation.

Table 2.3 Criteria for CALL task appropriateness (Chapelle, 2001 p. 55)

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-learning potential</td>
<td>The degree of opportunity present for beneficial focus on form.</td>
</tr>
<tr>
<td>Learner fit</td>
<td>The amount of opportunity for engagement with language under appropriate conditions give learner characteristics.</td>
</tr>
<tr>
<td>Meaning focus</td>
<td>The extent to which learners’ attention is directed toward the meaning of the language.</td>
</tr>
<tr>
<td>Authenticity</td>
<td>The degree of correspondence between the CALL activity and target language activities of interest to learners out of the classroom.</td>
</tr>
<tr>
<td>Positive impact</td>
<td>The positive effects of the CALL activity on those who participate in it.</td>
</tr>
<tr>
<td>Practicality</td>
<td>The adequacy of resources to support the use of the CALL activity.</td>
</tr>
</tbody>
</table>

Three criteria from Chapelle’s framework (2001)—language-learning potential, learner fit, and positive impact—are integrated into the current project’s logical framework for CALL corrective feedback evaluation. A detailed description of each component in the logical framework, coupled with evaluation criteria, is illustrated in Figure 2.6. At the first stage of the framework, the theoretical foundations (the first component, comprising the
Interactionist Approach with the Noticing and Output Hypotheses, and Skill-Learning Theory) direct the pedagogical design of the CALL corrective feedback. These approaches inform the five main characteristics of the pedagogical design of CALL corpus-based feedback (the second component): 1) error noticing, 2) understanding enhancement, 3) output-inducing, 4) declarative knowledge learning. The CALL corrective feedback is used to develop error correction exercises (the third component) with the intention of promoting students’ immediate error noticing and correction; students are expected to find the design of the feedback helpful for their error corrections (the fourth component). The assumed immediate changes, then, are expected to lead to learning gains in the long term (the fifth component): students are able to transfer knowledge from the feedback to error correction when revising a new article. Finally, such improvement is expected to foster students’ writing accuracy and self-regulation, which is the ultimate goal of learning with CALL corpus-based corrective feedback (the sixth component).

In this logical framework, there are three hypotheses concerning the learning effects and impact of CALL corrective feedback: the CALL corrective feedback is expected to instigate immediate changes in the learning process, facilitate learning gains in the long term, and culminate in the ultimate goal of learning. Three criteria from Chapelle’s (2001) framework are used to investigate those hypotheses. Specifically, the evaluation criteria for language-learning potential and learner fit require researchers to collect quantitative and qualitative evidence to investigate the two hypotheses concerning immediate changes and long-term learning gains. The third criterion, positive impact, is used to investigate the hypothesis concerning the ultimate goal of learning with CALL corrective feedback. The evaluation criterion for each hypothesis is listed in Figure 2.6.
Figure 2.6 Logical framework for the development and evaluation of CALL corrective feedback.
2.4.3. Research Questions

Corresponding to the three main hypotheses in the logical framework for evaluating CALL corrective feedback and their evaluation criteria, this study aimed to collect evidence with which to address the following research questions (seen in Table 2.4) to evaluate the effectiveness of corpus-based corrective feedback in improving syntactic knowledge learning in academic writing.

Table 2.4 Research questions and evaluation criteria in this study

<table>
<thead>
<tr>
<th>Quality</th>
<th>Evaluation Criteria</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language-learning potential</strong></td>
<td>Data indicating error noticing and correction in error correction exercises with CALL corrective feedback</td>
<td>1. What evidence suggests that students with corpus-based feedback are better equipped to notice syntactic errors in the learning process than students with coded feedback?</td>
</tr>
<tr>
<td></td>
<td>Data indicating error correction in the pre- and post-tests without any feedback</td>
<td>2. What evidence suggests that students with corpus-based feedback are better equipped to correct syntactic errors in the learning process than students with coded feedback?</td>
</tr>
<tr>
<td><strong>Learner fit</strong></td>
<td>Assessment of students' perceptions of the appropriateness of the CALL corrective feedback for error correction</td>
<td>3. What evidence suggests that students treated with corpus-based feedback are better equipped to apply syntactic knowledge to correct errors in a new text as an outcome of the learning than students treated with coded feedback?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. What evidence suggests that advanced-low ESL learners perceive that corpus-based feedback is more appropriate than coded feedback for enabling correction of syntactic errors?</td>
</tr>
</tbody>
</table>
Table 2.4 Continued

<table>
<thead>
<tr>
<th>Quality</th>
<th>Evaluation Criteria</th>
<th>Research Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive impact</td>
<td>Assessment of students’ perceptions of positive learning experience with CALL corrective feedback</td>
<td>5. What evidence suggests that students treated with corpus-based feedback have a more positive learning experience than those treated with coded feedback?</td>
</tr>
</tbody>
</table>

2.5. Chapter Summary

This chapter reviewed the literature that inspired the dissertation and articulated theoretical foundations for developing and evaluating corpus-based corrective feedback. First, the chapter discussed two unsettled issues concerning the effects of corrective feedback on error correction and knowledge learning and unveiled the research gaps for further investigation. Then the chapter investigated the direct and indirect use of learner corpus data in CALL tool development and examined the evaluation of the effectiveness of the corpus-based pedagogical materials and tools in the existing literature. This review of the literature on corrective feedback and CALL corpus-based tool research has provided an adequate justification for the research goals and research questions guiding the dissertation.

The chapter also delineated two theoretical approaches to SLA—the Interactionist Approach, and Skill Learning Theory—and considered the ways they have informed the development of corpus-based corrective feedback. The Logic Model of the Theory of Action (Patton, 2008) and Chapelle’s (2001) CALL task evaluation criteria were discussed to illustrate how they were integrated to develop an evaluation framework for CALL corrective feedback and to construct the research questions. The investigation of
theoretical approaches to SLA, the Logic Model of Theory of Action (Patton, 2008), and Chapelle’s (2001) evaluation criteria comprised a solid theoretical foundation for evaluation of corpus-based corrective feedback.
CHAPTER 3. METHODOLOGY

This chapter presents the methodology used in this dissertation. To address the five research questions concerning language learning potential, learner fit, and positive impact, the study employs a mixed-methods approach to collect quantitative and qualitative data from 90 participants to evaluate and compare the effects of corpus-based and coded feedback. The chapter begins with a description of the mixed-methods approach with its convergent design. Following that, it provides a detailed account of the study's participants and sampling methods, teaching context, and target error types, and of the materials and website used to collect quantitative and qualitative data. Moreover, the chapter describes the detailed data collection procedures undertaken inside and outside of the classroom. Finally, the analyses of quantitative and qualitative data for each research question are discussed.

3.1. Research Design

To address the five research questions, this study adopted a mixed-methods approach with a convergent design to investigate the respective effects of corpus-based feedback and coded feedback on students' error corrections exercises and learning. According to Creswell and Clark (2011), a mixed-methods approach combines quantitative and qualitative approaches to offer a better understanding of research problems. One of the main purposes of convergent design is to “illustrate quantitative results with qualitative findings, synthesizing complementary quantitative and qualitative results to develop a more complete understanding of a phenomenon, and comparing multiple levels within a system” (p. 77). Using convergent design, this study collected quantitative and qualitative
data separately and incorporated information from those two perspectives for data analysis and interpretation.

In this study, quantitative and qualitative data about students’ use and perceptions of target corrective feedback were collected separately during one phase of the research. As seen in Figure 2.7, the quantitative data consisted of the scores that measured the immediate effects of feedback in error correction exercises, the scores that measured learning gains demonstrated by the pre- and post-tests, and students’ ratings of the effects of the feedback in Likert-scale questionnaires. The qualitative data consisted of think-aloud exercises and semi-structured interviews, which played an important role in describing participants’ learning experiences with the intervention. Each set of data was analyzed independently to generate its own initial results. In the merging and interpreting stage, the content areas represented by both sets were identified and the results were synthesized and interpreted together to reach a well-substantiated conclusion.

Figure 3.1 The mixed-methods approach with convergent design applied in this study.
Employing a mixed-methods approach with a convergent design, this study compared the results from the data sets on both types of feedback. This comparison provided a well-substantiated conclusion concerning the effectiveness of the two types of feedback on students’ error correction and language development. The results also shed light on the relationship of each type of feedback to learner fit and on their overall impact on learning.

3.2. Participants and Sampling

This study was conducted at a large university in the Midwestern United States. The participants comprised 90 students enrolled in different sections of the writing course English 101C, Academic Writing II for International Undergraduates. All the students achieved the minimum scores on English proficiency tests required by the university (TOEFL iBT, 72 or IELTS 6.0). They also demonstrated their English writing proficiency as advanced-low on an institutional English Placement Test (EPT) based on EPT rating rubrics developed from the ACTFL (American Council on the Teaching of Foreign Languages) writing proficiency assessment framework, leading directly to their placement in English 101C. Students attended the course three hours per week and submitted four major assignments in one semester. Instructors used error codes—abbreviated labels—for different kinds of grammatical errors when marking student major assignments.

The demographic information of the participants, such as age, gender, nationality, native language, major and degree pursued, TOEFL or IELTS training, and duration of time spent living in the U.S. was obtained from a background survey and is described in Table 3.1. There were 45 participants in the code group and 45 in the corpus group. A majority of
participants spoke Chinese as their first language, while the rest spoken Korean, Malay, Arabic, French, Russian, and Thai. Participants’ ages were between 19 to 22 years. Their TOEFL iBT scores ranged between 72-99 and IELTS scores ranged between 6.0-6.5. Most of the participants came to the U.S. less than 6 months prior to the study. All the students were competent in computer and internet use.

Table 3.1. *Demographic information about participants*

<table>
<thead>
<tr>
<th></th>
<th>Code Group</th>
<th>Corpus Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of students</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Sex</td>
<td>Female (17)</td>
<td>Female (19)</td>
</tr>
<tr>
<td></td>
<td>Male (28)</td>
<td>Male (26)</td>
</tr>
<tr>
<td>Ages</td>
<td>19-22</td>
<td>18-22</td>
</tr>
<tr>
<td>L1</td>
<td>Chinese (29)</td>
<td>Chinese (28)</td>
</tr>
<tr>
<td></td>
<td>Korean (6)</td>
<td>Malay (6)</td>
</tr>
<tr>
<td></td>
<td>Arabic (5)</td>
<td>Korean (6)</td>
</tr>
<tr>
<td></td>
<td>Malay (2)</td>
<td>Vietnamese (2)</td>
</tr>
<tr>
<td></td>
<td>Thai (2)</td>
<td>Arabic (2)</td>
</tr>
<tr>
<td></td>
<td>Russian (1)</td>
<td>French (1)</td>
</tr>
<tr>
<td>TOEFL iBT or IELTS scores</td>
<td>iBT 73-99</td>
<td>iBT 72-97</td>
</tr>
<tr>
<td></td>
<td>IELTS 6.0-6.5</td>
<td>IELTS 6.0-6.5</td>
</tr>
<tr>
<td>Time in the U.S.</td>
<td>Less than 6 months (39)</td>
<td>Less than 6 months (37)</td>
</tr>
<tr>
<td></td>
<td>Between 6 months and 1 year (2)</td>
<td>Between 6 months and 1 year (4)</td>
</tr>
<tr>
<td></td>
<td>Between 1 and 2 years (1)</td>
<td>Between 1 and 2 years (3)</td>
</tr>
<tr>
<td></td>
<td>Between 2 and 4 years (3)</td>
<td>Between 2 and 4 years (1)</td>
</tr>
</tbody>
</table>

For the in-class activities (error correction practice exercises, pre- and post-tests, and questionnaire), half of the study participants in each section of the course were randomly assigned to the control group and half were assigned to the treatment group. Given that students in the control and the treatment groups were registered in the same
classes, the study was able to control for variables in terms of the impact of instructors on student learning.

Eighteen students, nine from each group, voluntarily participated in out-of-class research; they were scheduled to meet in a conference room individually for think-aloud protocols and semi-structured interviews after the post-test. To collect representative input from the two groups, those 18 students were selected based on their pre-to-post-test gain scores (arrived at by subtracting pre-test scores from post-test scores for each person) and were categorized into three improvement levels (high-, mid-, and low-improvement), as shown in Table 3.2. For each participant, the think-aloud protocols and the interview lasted approximately 40 minutes and were audio-recorded.

Table 3.2 Sampling design of selected participants in think-aloud protocols and semi-structured interviews

<table>
<thead>
<tr>
<th>Pre-to-post-test gain scores</th>
<th>Code group</th>
<th>Corpus group</th>
</tr>
</thead>
<tbody>
<tr>
<td>High improvement</td>
<td>Three participants</td>
<td>Three participants</td>
</tr>
<tr>
<td>Mid improvement</td>
<td>Three participants</td>
<td>Three participants</td>
</tr>
<tr>
<td>Low improvement</td>
<td>Three participants</td>
<td>Three participants</td>
</tr>
</tbody>
</table>

3.3 Teaching Context

English 101, Academic Writing II for International Undergraduates, included four units. The course taught students to perform description, analysis, critique, synthesis, and evaluation in different genres of academic papers and to proof, edit, and correct drafts for syntax, mechanics, and word choice. Students enrolled in English 101C attended the course three hours per week and submitted four major assignments over the course of the semester.
With one of main goals being to guide students in correcting errors, English 101C covered advanced points of English grammar as they apply to English academic writing. In each unit, students began by composing 400- to 600-word essays on a target topic and then receiving teacher feedback for a subsequent revision. Corrective feedback played an important role in addressing errors in context, enhancing students’ knowledge, and facilitating language learning. English 101C instructors were required to use coded feedback consisting of abbreviated labels for different kinds of grammatical errors when marking students’ writing. In the interest of improving the quality of corrective feedback and teacher-student communication, a corpus-based feedback system called My Feedback was developed, and teaching units addressing syntactic errors with corpus-based feedback were integrated into the English 101C syllabus in the spring and fall 2017 semesters.

3.3. The Focused Error Types of this Study and Their Definitions

This study focused on the effects of feedback on three types of syntactic errors: sentence fragments, run-on sentences, and subordinate clause errors. These errors were not taught explicitly in English 101C before the study; however, instructors marked or corrected various grammatical errors, including those three types of syntactic errors, in students’ major papers.

Various definitions of sentence fragments (Ferris, 2014; Harris, 1981; Kline & Memering, 1977) have appeared in the literature. Kline and Memering (1977) defined the sentence fragment as “a broken sentence” and introduced the term “incomplete thoughts/ideas” to describe one of the key features of fragments. According to Shaughnessy (1977), the term “a complete idea” in this definition is itself very vague. He (1977) explained that some sentence fragments, especially dependent structures, include a
“complete idea” and have a sense of rhetorical independence in a specific context. Therefore, it is hard to establish criteria to distinguish sentence fragments from non-fragments. The most recent definition, from Ferris (2014), states that a sentence fragment may be missing either a subject or a verb, or it may consist solely of one or more dependent clauses. This definition is clearer because it attempts to discuss the incompleteness of a sentence from a syntactic rather than a rhetorical perspective. However, it is not clear whether the missing subject or verb is located in the main clause or the subordinate clause. Since none of the existing definitions prove to be comprehensive, this study defines a sentence fragment as a construction that may exhibit any of the characteristics in Table 3.3.

Run-on sentences are another type of syntactic error that the dissertation focused on. Nielsen (1984) discussed two types of run-on sentences: 1) a fused sentence, or two independent clauses joined without punctuation, and 2) a comma splice, or two independent clauses linked by only a comma. Considering that a run-on sentence may include more than two independent clauses, Ferris (2014) redefined run-on sentences: “When two or more sentences are run together without appropriate punctuation or other connectors, ... the result is a run-on sentence” (p. 339). A comma splice error is a specific type of run-on sentence in which two or more sentences are joined by a comma. In this study, both fused sentences and comma splice errors are categorized as run-on sentences; the definition is provided in Table 3.3.

The dissertation explored the effects of corrective feedback on students’ ability to correct errors at the clause level in texts that were similar to those that they might write in the English 101 course. Sentence structure errors are ranked as the most frequent errors in
ESL writing, accounting for more than 20% of the total errors identified by Ferris et al. (2000, cited in Ferris, 2002). Clause errors are one type of sentence structure error. They are discussed in many ESL textbooks (Fitzpatrick, 2011; Lane & Lange, 2011). For example, Lane and Lange (2011) defined a clause error as “an error in which the formation of a relative, adverbial or noun clause is incorrect” (p. 124). This definition is used in this study. The possible features of clause errors are discussed in Table 3.3.

Table 3.3 *Definitions and main features of the three syntactic errors in this study*

<table>
<thead>
<tr>
<th>Sentence fragments</th>
<th>Run-on sentences</th>
<th>Clause errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Lacking a subject, a main verb, or both in the main clause; or</td>
<td>1) A fused sentence: two or more independent clauses joined without punctuation; or</td>
<td>1) Improper coordination, subordination, or conjunction between clauses; or</td>
</tr>
<tr>
<td>2) containing one or more dependent clauses without an independent clause attached; or</td>
<td>2) a comma splice: two or more independent clauses linked only by a comma.</td>
<td>2) incorrect word order; or</td>
</tr>
<tr>
<td>3) only consisting of one phrase, or multiple phrases (noun, preposition, adverb, conjunction phrases, etc.).</td>
<td></td>
<td>3) inappropriately omitted words or phrases, or unnecessary words or phrases in the clauses.</td>
</tr>
</tbody>
</table>

### 3.3 Materials

This section delineates the materials used in this study, including the *My Feedback* website, the error correction practice exercises, the error awareness sheet, the pre- and post-tests, the Likert-scale questionnaires, the think-aloud protocols, and the semi-structured interview questions.
3.3.1. My Feedback Website

*My Feedback* is a web-based corrective feedback system to which students are able to submit essay drafts and receive corpus-based feedback (Figure 3.2). The system includes teacher and student interfaces.

![Welcome to My Feedback](image)

**Figure 3.2 Screenshot of My Feedback website.**

Using the teacher interface, instructors are able to access a corpus-based feedback library. The construction of error categories and their patterns in this library was informed by an EPT learner corpus from a large university in the Midwestern United States. This library currently stores corpus-based feedback on three error categories: sentence fragments, run-ons, and subordinate clause errors. When instructors locate an error in an essay, they right-click the error to access the feedback library, choose corresponding
feedback, and insert it into the document. Teachers can also enter any further feedback into a text editor to address grammar, content, and structure issues in students’ writing.

On the student webpage, students can type their essays into a text editor directly or copy-and-paste their pre-written drafts into the system. When students receive feedback from their instructors, they find the erroneous sentences marked with a red “X.” If students hover the cursor above the X, corrective feedback related to the target error pops up in a new window and students can correct the error accordingly. By clicking the “submit” button, students submit their revised version for grading.

I developed the My Feedback website for the study and for future use in ESL writing classrooms due to the limited number of commercial and research-based websites that allow instructors to provide corrective feedback to students. Table 3.4 contains descriptions of existing commercial websites and programs that provide corrective feedback.

Table 3.4 Descriptions of websites and programs providing corrective feedback

<table>
<thead>
<tr>
<th>Website /Program</th>
<th>Website type</th>
<th>Description</th>
<th>Feedback types</th>
<th>Features of feedback system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaizena (2018)</td>
<td>Commercial website and app</td>
<td>Kaizena is an online program that assists teachers in providing voice comments and personalized feedback on student work. Kaizena allows instructors to save frequent comments so that they can reuse them conveniently.</td>
<td>1. Written feedback  2. Audio feedback  3. Video lessons/feedback</td>
<td>Kaizena allows instructors to highlight erroneous parts of the sentence and embed video feedback into the assignment.</td>
</tr>
</tbody>
</table>
Table 3.4 Continued

<table>
<thead>
<tr>
<th>Website/Program</th>
<th>Website type</th>
<th>Description</th>
<th>Feedback types</th>
<th>Features of feedback system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Doc</td>
<td>Commercial</td>
<td>Google Doc/Microsoft is a word processor website/program where instructors can provide written feedback to students.</td>
<td>1. Written feedback</td>
<td>Google Doc includes the unique feature “Revision History,” a catalog of document changes. This feature allows instructors to see the changes made to revised drafts.</td>
</tr>
<tr>
<td>Mark My Words</td>
<td>Research-based</td>
<td>Informed by corpus analysis, Mark My Words software includes brief pre-written comments that can be inserted in students’ papers to address repeated lexical and stylistic errors.</td>
<td>3. Links to outside websites</td>
<td>Mark My Words includes preset feedback on lexico-grammatical errors. The instructor can also provide individualized feedback.</td>
</tr>
<tr>
<td>(2010)</td>
<td>feedback website</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As described in Table 3.4, the feedback systems in existing commercial and research-based websites and programs include the following crucial features that facilitate teacher-student communication and teacher feedback:

- All the websites and programs enable instructors to offer customizable written feedback to language learners.

- Kaizena (2018) and Google Doc (2018)/Microsoft Word (with add-on) (2016) offer an option for audio feedback. Instructors can record their voices and insert them as audio feedback into specific locations in students’ writing; Kaizena (2018) also supports the use of video feedback and video lessons.
• Kaizena (2018) and Mark My Words (2010) include a comment library. In Kaizena, instructors can produce their own written, audio, or video feedback and save all types of feedback in a comment library for future use. In Mark My Words (2010), researchers have already included over a thousand pages of comprehensive metalinguistic feedback and links to writing-related information in the system. Instructors can either extract comments from the library or develop their own written feedback to insert into students’ work.

• Kaizena (2018) also supports the integration of video feedback or video lessons into written assignments. Google Doc (2018)/Microsoft Word (2016) and Mark My Words (2010) allow instructors to include hyperlinks in marginal comments.

Even though the websites mentioned above support the provision of corrective feedback, I have chosen to design my own feedback system instead of using existing websites or programs. Table 3.5 compares the main features of the existing feedback systems or apps to My Feedback system.

Table 3.5 Corrective feedback programs/apps and their features

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In-text written comments</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Oral comments</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Video comments/lessons</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 3.5 *Continued*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Image-format comments</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Comments sharing among instructors</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Comment library</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pre-set comment</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Compared to other programs, *My Feedback* includes more features to facilitate feedback provision. First, the written corpus-based feedback in *My Feedback* is saved in an image-file format in the feedback library; as seen in Table 3.5, none of the existing commercial and research-based websites allow for image-format feedback. The choice to use image-format feedback was based on students’ user experience. When designing the computer-assisted corrective feedback, I had originally chosen to create video feedback. I recruited five students from English 101C to use video feedback for error correction. Surprisingly, those students found the pace of the videos to be either too slow or too fast for their purposes. What was more, they pointed out that simultaneously reading written feedback and listening to audio explanations was overwhelming as they attempted to assimilate the information contained therein; moreover, they mentioned that the buffering times and advertisements associated with video streaming were distracting and annoying. Based on students’ suggestions, I designed image-format feedback and decided to build my own website on which to save those files.
Another reason behind my decision to construct the *My Feedback* website was my desire to create a corrective feedback library that may be used by English 101C instructors in the future. In English writing courses, instructors’ ability to provide high-quality explicit feedback is always restricted by time limitations and class sizes. At the same time, instructors spend a great deal of time composing comments to address recurrent errors in students’ writing. One way to improve the efficiency of providing teacher feedback is to compile a comment library (Milton, 2006). The existing feedback websites and programs do not offer opportunities for instructors to create and share corrective feedback for paper editing. With the *My Feedback* system, instructors who encounter repeated error patterns can respond with preset feedback from the library; they can also design their own feedback and share it with others in the feedback library. With these principles in mind, I worked with my colleagues in the Department of Computer Science to design the *My Feedback* website.

In this study, students in the corpus group received corpus-based feedback from *My Feedback* on their error correction practice exercises, consisting of a text with errors marked that the students are asked to correct. The corpus-based feedback provided by *My Feedback* exhibits four main features: it is 1) exemplar-based, 2) color-coded, 3) metalinguistic, and 4) audiovisual, as seen in Figure 3.3.
Figure 3.3 Screenshot of a sample of corpus-based feedback in My Feedback

First, the feedback is exemplar-based. Each piece of feedback includes an incorrect sentence containing a similar error and its correction to illustrate the target error pattern and options for revision. The incorrect sample sentences were extracted by the primary researcher and an undergraduate student majoring in English from an EPT corpus and modified to exemplify typical patterns of sentence fragments, run-ons, or subordinate clause errors. The corresponding corrected sentences are used to illustrate the methods for correction. Second, the feedback is color-coded: red is used to mark the error and green is used to indicate the required correction. Third, the feedback includes a metalinguistic explanation in which metalinguistic terms (such as dependent clauses in Figure 3.3) are attached to the sample sentences or used to help explain the structure of the sentences and the nature of the target errors. Fourth, an audio explanation is available as an option to accompany the written feedback.

In this study, the coded feedback was provided to students via My Feedback. Providing coded feedback and corpus-based feedback through the same system helped eliminate confounding variables when comparing the two types of feedback. Coded feedback uses written symbols (e.g. frag for sentence fragments; ro for run-on sentences; cl
for clause errors) to indicate error categories and facilitate students’ error correction. Figure 3.4 shows an example of coded feedback for a sentence fragment.

![Figure 3.4 Screenshot of a sample of coded feedback.](image)

**3.3.2. Error Correction Exercises in My Feedback**

Two computer-based error correction exercises were developed to evaluate the immediate effects of corpus-based and coded feedback on students’ error corrections. Participants were provided with either corpus-based or coded feedback to correct sentence fragments, run-on sentences, and subordinate clause errors in these two exercises.

Each error correction exercise required participants to use the target feedback they receive to correct errors in a 600-word article. Participants were familiar with the topics of the articles because they had been asked to compose an essay on the same prompt for a major assignment before beginning the error correction exercises. The article used in the first exercise discussed Korean culture from three perspectives; the article in the second exercise focused on three Malaysian Chinese cultures and their related customs. Both articles were written by previous 101C students and were revised by the primary researcher, and the second researcher, an undergraduate student majoring in English for this study. Each article included 21 error patterns: 7 sentence fragments, 7 run-on sentences, and 7 clause errors. Although the content in the two articles was different, the
error patterns were the same, as was the feedback provided for each pattern. To access the exercises, participants logged into their individual accounts on the My Feedback website, as seen in Figure 3.5. In each exercise, the article, placed on the left side of the screen, was non-editable. Certain sentences in the article were flagged with a red X; students were informed that each of these marked sentences included one grammatical error. When the cursor hovered over a red X, a feedback window popped up. An editable version of the same article appeared in the text editor on the right side of the screen; students used the target feedback to correct errors and saved their changes by clicking the “submit” button, as seen in Figure 3.5.

![Figure 3.5 Screenshot of error correction exercise 1 in My Feedback.](image-url)
3.3.3. Error Awareness Sheet

The Error awareness sheet was distributed as a document to the group receiving coded feedback to serve as supplementary material for the error correction practice exercises. Coded feedback is the current form of feedback used to mark grammatical errors in students’ major writing assignments in English 101C. When 101C students received coded feedback, they were encouraged to consult the error awareness sheet (Appendix A) to look up the error codes. Similarly, when participants in the code group completed the error correction practice exercises, they consulted the error awareness sheet for the meanings of the codes or for a brief description of the errors.

3.3.4. Pre- and Post-tests

To evaluate students’ knowledge of sentence fragments, run-ons, and subordinate clause errors before and after their use of the feedback, a pre- and post-test design was implemented. The pre-test was paper-based and comprised one task: students were required to read a 600-word article, finding and correcting errors in it within 35 minutes. The article was separate from those used in the two error correction exercises. Without receiving any feedback, students needed to use their prior knowledge to identify and correct errors in the pre-test. A post-test was administered two weeks after the second error correction exercise. The texts used in the pre- and post-tests and the requirements of the tests were the same for two reasons. First, the use of the same article ensured the consistency of the difficulty level of both tests. In other words, the sentence lengths, the vocabulary level, and other features of the texts were controlled variables. Gain scores in the pre- and post-test design therefore reflected students’ learning outcomes with the target feedback. Second, there was a four-week interval between the pre- and post-tests.
Although some students might remember some of the content of the pre-test four weeks later, this recollection alone would not be sufficient to give them clues on error corrections on the post-test. The pre- and post-test sheets are presented in Appendix B.

### 3.3.5. Background Survey Questions

Before the pre-tests, students were given seven survey questions to elicit basic information, including their age, gender, first language, level of study, colleges attended, TOEFL or IELTS scores, and their time of study in the U.S. The background survey is included in Appendix C.

### 3.3.6. Likert-scale Questionnaire

Twenty-one 7-point Likert-scale items were used to measure learner fit and learning impact (see Appendix D). According to Weijters, Cabooter, and Schillewaert (2010), 7-point rating scales are appropriate for student populations. Therefore, a 7-point Likert-scale questionnaire, with points ranging from “Strongly Disagree” to “Strongly Agree,” was developed. The three sub-constructs of learner fit (clarity of feedback, usefulness of the feedback, and students’ learning strategies with the feedback) were evaluated through nine items. The Cronbach’s alpha reliability of those nine Likert-scale items was .862, indicating that they largely measured the same construct. The four sub-constructs of learning impact (satisfaction of the learning experience with the feedback, enthusiasm for using target feedback to correct other errors, error awareness raising with the feedback, and motivation stimulating with the feedback) were evaluated with twelve items. The Cronbach’s alpha reliability of those twelve Likert-scale items was .954, indicating a high level of internal consistency.
Three different items have been developed to measure each sub-construct. These three statements were rephrased in a number of ways: some of them used synonymy, paraphrasing, or different words from same word family to prevent the data from being distorted if the respondents misread the statements (e.g. items 1 and 2 in Figure 12); some of them were negatively worded to evaluate whether students read the questionnaire carefully and whether answers to those items were consistent (e.g. item 3 in Figure 12). Furthermore, all statements were carefully worded to avoid ambiguity and misinterpretation.

--- About levels of difficulties for understanding feedback

1. The feedback was **understandable in the error correction practices**.
   Strongly Disagree 1------- 2 ------- 3------- 4 ------- 5 ------- 6 ------- 7 Strongly Agree

2. The feedback **helped to explain** the errors clearly.
   Strongly Disagree 1------- 2 ------- 3------- 4 ------- 5 ------- 6 ------- 7 Strongly Agree

3. I had **difficulties** in understanding the feedback.
   Strongly Disagree 1------- 2 ------- 3------- 4 ------- 5 ------- 6 ------- 7 Strongly Agree

Figure 3.6 *Three items about the comprehensibility of the feedback in the questionnaire.*

Three different items concerning each construct have been developed to increase construct validity of the questionnaire and the reliability of the results. Student responses were screened for abnormalities by checking the consistency of learners’ responses to the items for each sub-construct assessed by the questionnaire.

### 3.3.7. Instructions for the Think-aloud Protocols

Instructions for the think-aloud protocols (see Appendix E) were provided to nine students from each group who had volunteered to participate in a think-aloud activity outside of class. Those 18 students were asked to revisit the error correction practice
exercises. Students were asked to describe their observations, thinking processes, feelings, and concerns in real time as they interacted with the feedback in the error correction exercises to notice and correct nine syntactic errors. A five-minute demonstration was performed by the primary researcher before the activity, and questions from students were addressed. During the think-aloud protocol, the researcher encouraged students to verbalize their ideas when they have difficulties, and asked questions when needed to clarify students' answers. Each participant's response was recorded with Audacity for data analysis and interpretation.

3.3.8. Semi-structured Interview Questions

On a volunteer basis, the same 18 students who attended the think-aloud protocols participated in the 20-minute semi-structured interviews. Based on students' performance on the error correction exercises, they were asked to evaluate the effects of the feedback on noticing and correcting sentence fragments, run-ons and subordinate clause errors respectively. According to their pre- and post-test results, they were also asked to explain reasons behind their successful and unsuccessful error correction behaviors in the post-test. In addition, students would articulate the appropriateness of the feedback for their language learning, and the impact of learning. The interviews provided insight into students' subjective experiences of the program, shedding light on the quality of the feedback and on students' performance and language learning.
3.4. Procedure

With the support of the English Department and the English 101C coordinator, I coordinated with the course instructors to integrate the research activities for the study into the second unit of the English 101C course. A pre- and post-test design was implemented to evaluate the effects of corpus-based feedback and coded feedback on students’ error correction and language learning. Prior to the data collection, this study was approved by the Institutional Review Board (IRB) (Appendix F). The data collection included in-class activities and out-of-class research. The in-class activities comprised a demographic survey, a pre- and post-test, two error correction exercises, and a questionnaire; the out-of-class part comprised think-aloud protocols and semi-structured interviews. The data collection procedure is illustrated in Figure 3.7.

![Procedure Diagram]

Figure 3.7 The procedure of this study.
3.4.1. In-class Activities

In week 1 of this project, with the permission of the course coordinator and instructors, the primary researcher visited eight sections of English 101C. Study participants completed a demographic survey and wrote a pre-test lasting 35 minutes. Prior to the pre-test, students were given examples of grammatically flawed sentences and instructions on the appropriate error correction procedures.

In week 2, all participants worked on the first error correction exercise in their regular computer lab classroom. Half of the participants in each class were randomly assigned to the code group receiving coded feedback, and the second half were assigned to the corpus group receiving corpus-based feedback. Each participant was given a username and a password to log into the My Feedback website and was asked to use target feedback to correct 21 errors in a 600-word article within 40 minutes. A 5-minute demonstration on how to use feedback to revise errors on the webpage was given before the exercise. In the regular English 101C class, when students received coded feedback from their instructor, they could access a referential document called the error awareness sheet (Appendix E) to check the error codes. In this error correction exercise, students in the contrast group could refer to that document for the meanings of the codes or a brief description of the errors. Students in the corpus group could only use corpus-based feedback to correct errors.

In week 3, students remained in the same groups to complete the second error correction exercise. The requirements of this exercise were the same as the previous one: participants were required to use either coded or corpus-based feedback to correct 21
errors in a new 600-word article within 40 minutes. The error patterns and the accompanying feedback in both error correction exercises were the same, but the article was different. The repeated exposure to the same error patterns and feedback was meant to enforce students' knowledge of the target errors and corrections.

In week 5, two weeks after the second error correction exercise, the primary researcher met students again in a standard classroom to administer a paper-based post-test. This test was administered in the same manner as the pre-test: participants were required to identify and correct errors in a 600-word article within 35 minutes. After finishing the post-test, participants completed a questionnaire about their learning experience with the target feedback.

3.4.2. Out-of-class Research

Students who voluntarily joined the out-of-class activities participated in a think-aloud protocol activity and a semi-structured interview. Each individual was scheduled to meet with the primary researcher in a conference room. During the think-aloud protocols, students interacted with the corrective feedback in *My Feedback* again and articulated their feelings and thoughts for 20 minutes. After the think-aloud activity, they were asked follow-up questions in a 20-minute semi-structured interview in which they further described their learning experiences with the feedback. These two activities were be audio-recorded for coding and data analysis.
3.5. Data Analysis

Table 3.7 summarizes the qualitative and quantitative data analysis methods for each of the dissertation’s research questions.

<table>
<thead>
<tr>
<th>Evaluation Criterion</th>
<th>Research Focus</th>
<th>Data</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Learning Potential</td>
<td>1. The effects of each type of feedback on error noticing in the learning process</td>
<td>Quan: Error correction exercises Qual: Think-aloud activity; semi-structured interview</td>
<td>Descriptive analysis and one-way ANOVAs Thematic analysis</td>
</tr>
<tr>
<td></td>
<td>2. The effects of each type of feedback on error correction in the learning process</td>
<td>Quan: Error correction exercises Qual: Think-aloud activity; semi-structured interview</td>
<td>Descriptive analysis and one-way ANOVAs Thematic analysis</td>
</tr>
<tr>
<td></td>
<td>3. The effects of each type of feedback on knowledge application following the learning process</td>
<td>Quan: Scores from pre- and post-tests Qual: Think-aloud activity; semi-structured interview</td>
<td>Descriptive analysis and two-way mixed ANOVA Thematic analysis</td>
</tr>
<tr>
<td>Learner Fit</td>
<td>4. The appropriateness of each type of feedback for error correction</td>
<td>Quan: Questionnaire Qual: Think-aloud activity; semi-structured interview</td>
<td>Descriptive analysis and one-way ANOVA Thematic analysis</td>
</tr>
<tr>
<td>Positive Impact</td>
<td>5. The impact of each type of feedback on language learning</td>
<td>Quan: Questionnaire Qual: Think-aloud activity; semi-structured interview</td>
<td>Descriptive analysis and one-way ANOVA Thematic analysis</td>
</tr>
</tbody>
</table>
Research Question 1: What evidence suggests that students with corpus-based feedback are better equipped to notice syntactic errors in the learning process than students with coded feedback?

The first research question was addressed by analyzing students’ error noticing scores on two error correction practice exercises and coding the transcripts of students’ think-aloud protocols and semi-structured interviews.

Each error correction practice exercise included 21 syntactic errors in three categories: sentence fragments, run-ons, and subordinate clause errors. When a student’s correction demonstrated his or her attention to all erroneous elements of a syntactic error, the student received 1 point for that item. When a student’s change revealed attention only to part of the total possible erroneous elements, the student received 0.5 points for the item. Finally, when a student failed to demonstrate any recognition of the error, 0 points were awarded for that item. Each student received a holistic score per exercise, and all exercises were fully coded by the primary researcher. A second researcher recoded 20% of the data (18 exercises total—9 from the corpus group and 9 from the coded group) in each error correction exercise. Cohen’s Kappa inter-rater reliability coefficient between the two researchers was .881 in error correction practice one and .877 in error correction practice two.

In terms of quantitative data analysis, a one-way ANOVA was used to compare the total scores on each error correction practice exercise between the two groups. Assumptions of outliers, normality, and homogeneity of variances were tested. With the data from error correction practice exercise one, the assumption of homogeneity of
variances was not met and Welch’s ANOVA was reported since this statistic is appropriate when the homogeneity of variance assumption is not met (Brown & Forsthe, 1974; Field, 2013). Welch’s ANOVA was reported for the data from error correction practice two, because the assumption of homogeneity of variances was violated for these data as well, The descriptive data of students’ noticing scores for the three types of syntactic errors in each error correction exercise were also reported.

In terms of qualitative analysis, think-aloud protocols and semi-structured interviews were conducted to gather in-depth information about students’ experiences in error noticing with the target corrective feedback. During the think-aloud protocols, students revisited the error correction practice exercises; they articulated their cognitive and dynamic processes while using target feedback to detect errors. In the semi-structured interview, students discussed their experience of using the target feedback to notice each error type.

Two undergraduate students majoring in English were recruited to transcribe qualitative data gathered from the 18 participants. All transcriptions were thoroughly reviewed by the primary researcher to ensure their descriptive validity. Data analysis started with two researchers undertaking deductive coding of each transcript. The primary and second researchers first worked separately on the open coding for two interview transcripts on students’ experience noticing errors: two people assigned codes to the data in the margins of the transcriptions. Next, both researchers assessed and discussed their assignments together before reaching a consensus on code use in the data analysis. The comparison of coded data and discussion of coding differences helped to refine the overall
coding scheme. The refined coding scheme, with code use, code descriptions, and relevant examples, is delineated in Table 3.8.

Table 3.8 *Code use, code description, and examples related to error noticing with target feedback.*

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colors</td>
<td>Students’ thinking process and evaluation of the use of colors in the corpus-based feedback for error noticing</td>
<td>The red colors shows me the sentence is wrong and where is wrong. (Student 7, high-improvement learner; comment on a run-on sentence, corpus group)</td>
</tr>
<tr>
<td>Examples</td>
<td>Students’ thinking process and evaluation of the use of examples in the corpus-based feedback for error noticing</td>
<td>I read the sentences. It tells me the comma is not correct because there are two independent clauses. (Student 1, low-improvement learner; comment on a run-on sentence, corpus group)</td>
</tr>
<tr>
<td>Oral explanations</td>
<td>Students’ thinking process and evaluation of the use of oral explanation in the corpus-based feedback for error noticing</td>
<td>I’m an aural learner, so listening to the feedback is more efficient to understand where the error is and how to correct it. (Student 6, intermediate-improvement learner; comment on a subordinate clause error, corpus group)</td>
</tr>
<tr>
<td>Error codes</td>
<td>Students’ thinking process and evaluation of the use of error codes in the coded feedback for error noticing</td>
<td>The feedback says it’s a “FRAG.” I found the sentence doesn’t have a verb, so should add one. (Student 16, high-improvement learner; comment on a sentence fragment, code group)</td>
</tr>
</tbody>
</table>
Table 3.8 continued

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticing confirmation</td>
<td>Students’ thinking process and evaluation of corpus-based or coded feedback in confirming their identification of errors</td>
<td>I looked at the sentences, and search for errors. And I know “there are” is wrong and should be deleted. I looked at the feedback it tells me the same thing. (Student 5, intermediate-improvement learner; comment on a subordinate clause error, corpus group)</td>
</tr>
</tbody>
</table>

The primary and second researchers utilized the information in Table 3.8 to code idea units in four transcripts with Cohen’s Kappa reliability .916. The primary researcher then coded the remaining transcripts from the think-aloud protocols. Those codes were then used to construct two themes on error noticing: error noticing with feedback (colors, examples, oral explanations, and error codes), and error noticing confirmation with feedback (noticing confirmation). To evaluate students’ perceptions of error noticing with the two types of corrective feedback, the percentages of positive and negative idea units under those two themes were calculated and compared.

Students’ perceptions of the target feedback on each type of syntactic error from the semi-structured interviews were also analyzed. The researchers first worked on four transcripts individually and labeled each idea unit on sentence fragments, run-on sentences, and clause errors as either a positive or negative/partially negative comment. The inter-rater reliability of the two researchers was calculated to be 1.00. Then, 14 semi-structured interview transcripts were coded by the primary researcher. Based on the data analysis, the number of students who perceived the target feedback as helpful in alerting them to each error type was counted and compared between the two groups. Additionally,
students’ utterances about their perceptions of the usefulness of target feedback on each error were extracted to provide insight into their opinions about error detection with the target feedback on individual types of errors.

*Research Question 2: What evidence suggests that students with corpus-based feedback are better equipped to correct syntactic errors in the learning process than students with coded feedback?*

To answer the second research question, students’ error correction scores on the two error correction exercises, their responses on think-aloud activities, and their responses in the semi-structured interviews were collected and analyzed. Each error correction exercise included 21 syntactic errors. When students fully corrected a target error in the exercise without changing the original meaning of the sentence, they received 1 point for that item. When students corrected an error but changed the sentence’s original meaning, or if students only partially corrected an error, they received 0.5 points for that item. When students applied an incorrect revision or failed to attempt any correction for a given error, 0 points were awarded for that item. Each student received a holistic score per exercise. The two exercises were fully coded by the primary researcher and 20% of the data (18 exercises total—9 from the corpus group and 9 from the coded group) were coded by the second researcher. Cohen’s Kappa reliability between both researchers was .847 in error correction practice one and .867 in error correction practice two.

For quantitative data analysis, a comparison between the scores on each error correction practice exercise for the coded group and the corpus group was conducted using a one-way ANOVA with assumptions of outliers, normality, and homogeneity of variances being met. When the assumption of homogeneity of variance was violated (p = .00 < .05),
Welch’s ANOVA test was used to report the results. Descriptive data on both groups’ scores of correcting sentence fragments, run-ons, and subordinate clause errors were compared.

For qualitative data analysis, the think-aloud protocol and semi-structured interview were coded to extract idea units for theme construction. The data transcription and coding processes pertaining to error correction data were the same as the processes described in research question one. The coding scheme, with code use, code description, and corresponding examples about error correction, are described in Table 3.9.

Table 3.9 Code use, code description, and examples related to error correction with target feedback.

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples for error correction</td>
<td>Students’ thinking process and evaluation of the use of examples in corpus-based feedback for error correction</td>
<td>I look at the examples and know that I need to delete “in” because “in the book” is a preposition phrase, not a subject. So I also corrected the sentence. (Student 1, low-improvement learner; comment on a run-on sentence, corpus group)</td>
</tr>
<tr>
<td>Rules section for error correction</td>
<td>Students’ thinking process and evaluation of the use of the rules section in corpus-based feedback for error correction</td>
<td>The rule part shows the embedded questions should not put “did” before “he,” but it is not a question. So I deleted “does” and change. (Student 7, high-improvement learner; comment on a subordinate clause error, corpus group)</td>
</tr>
<tr>
<td>Oral explanations for error correction</td>
<td>Students’ thinking process and evaluation of the use of oral explanation in corpus-based feedback for error correction</td>
<td>1. The audio feedback explained it needs connectors between sentences. I read the example again and know the ways of correction. (Student 4, intermediate-improvement learner; comment on a run-on sentence, corpus group)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Basically, the oral feedback repeated what’s in the written feedback. I feel it’s not useful. (Student 2, low-improvement learner; comment on errors in general, corpus group)</td>
</tr>
</tbody>
</table>
The codes described in Table 3.9 were employed to code idea units in four transcripts with Cohen's Kappa reliability .833. The primary researcher then coded the remaining transcripts from the think-aloud protocols. Those codes were then used to construct two themes on error correction: error correction with feedback (examples, rules section, oral explanations, grammatical terms, general written feedback for error correction).
correction, and error codes for error correction) and error correction confirmation with feedback (correction confirmation). To evaluate students’ perceptions of error correction with the two types of corrective feedback, the percentage of positive and negative idea units under those two themes were calculated and compared, and sample utterances were elicited and discussed.

Qualitative data about students’ perceptions of the target feedback on correcting each type of syntactic error from the semi-structured interview were also analyzed. The coding process was the same as coding the semi-structured interview on error correction. Both the primary and second researchers individually coded two transcripts from the corpus group and two from the coded group, then labeled each idea unit from a given error category as positive or negative/partially negative. The inter-rater reliability calculated by Cohen’s Kappa was 1.00. Next, the primary researcher coded 14 semi-structured interview transcripts. Based on the data analysis, the number of students from each group who perceived the target feedback as helpful for correcting each error type was counted and compared. Students’ utterances on each error category were extracted to further illustrate their opinions.

Research Question 3: What evidence suggests that students treated with corpus-based feedback are better equipped to apply syntactic knowledge to correct errors in a new text than students treated with coded feedback?

Research question 3 was addressed by analyzing scores from pre- and post-tests and by coding students’ think-aloud protocols and semi-structured interviews. Twenty-one syntactic errors (7 sentence fragments, 7 run-ons, and 7 clause errors) were included in both the pre- and post-test, and the full score for each test was 21 points. When students
corrected an error and maintained the original meaning of the sentence, 1 point was awarded. When students corrected part of an error, or when they corrected an error but changed the original meaning of the sentence, 0.5 points were awarded. When the revision was incorrect or no attempt to correct the error was given, students received 0 points. The pre- and post-tests from the two groups were coded by the primary researcher. The second researcher recoded 20% of the data (18 pre-tests and 18 post-tests). Cohen’s Kappa reliability between both researchers was .843 in the pre-test and .815 in the post-test.

The descriptive statistics of students’ pre-test, post-test, and pre-to-post-test gain scores were reported. To examine the improvement from the pre-test to the post-test, a two-way repeated mixed-measures ANOVA test was conducted. The assumptions of outliers, normality, homogeneity of variance, and covariance were tested. When the assumption of outliers was violated, results with the original data including outliers and excluding outliers were reported and compared. Descriptive statistics of students’ pre-to-post-test gain scores on sentence fragments, run-ons, and subordinate clause errors were also reported.

Qualitative data from the think-aloud protocols about students’ successful and unsuccessful error correction in the post-test were coded and analyzed. The data transcription and coding processes about the qualitative data concerning knowledge learning were the same as processes described in research question one. The coding scheme, with code use, code description, and corresponding examples, was described in Table 3.10.
Table 3.10. Code use, code description, and examples concerning reasons for successful and unsuccessful error correction in the post-test.

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning from feedback</td>
<td>Successful error correction in the post-test due to learning from target feedback</td>
<td>So yeah, the exercises and the feedback, like examples, were really helpful to me. (Student 18, high-improvement learner; comment on a subordinate clause error, code group)</td>
</tr>
<tr>
<td>Pre-existing knowledge</td>
<td>Successful error correction in the post-test due to pre-existing knowledge (skipping the error in the pre-test)</td>
<td>I know how to correct it before. I think I just missed it [in the pre-test]. (Student 8, high-improvement learner; comment on a subordinate clause error, corpus group)</td>
</tr>
<tr>
<td>Personal practice</td>
<td>Successful error correction in the post-test due to learning from target feedback and practicing learned knowledge in one’s own writing</td>
<td>I guess the practice makes me improve. When I learn something, I want to use it in my life and then I learn gradually. (Student 18, high-improvement learner; comment on a subordinate clause, code group)</td>
</tr>
<tr>
<td>Not noticing errors without feedback or hints</td>
<td>Unsuccessful error correction in the post-test due to failed recognition of errors</td>
<td>If this sentence is marked as incorrect, I can find it and correct... like deleting “in” here, but if it’s not marked, I think it’s right. [Researcher: Why?] It sounds correct, and I just leave it there. (Student 5, intermediate learner; comment on sentence fragments, corpus group)</td>
</tr>
<tr>
<td>Forgetting</td>
<td>Unsuccessful error correction in the post-test due to forgetting what had been learned from the feedback.</td>
<td>It has been for a while, for a couple of days. I did not listen or rerevise based on the example. So I forgot (Student 9, intermediate-improvement learner; comment on a subordinate clause error, corpus group)</td>
</tr>
<tr>
<td>Code use</td>
<td>Code description</td>
<td>Examples from transcripts</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Complex sentences</td>
<td>Unsuccessful error correction in the post-test due to the complexity of sentences in the test</td>
<td>The sentence is kind of complicated, so I don't know how to deal with it. (Student 12, low-improvement learner; comment on a subordinate clause, code group)</td>
</tr>
<tr>
<td>Long sentences</td>
<td>Unsuccessful error correction in the post-test due to the length of sentences in the test</td>
<td>It's one of those sentences where you know it's a long sentence, and [you could] break it, but it's very hard. Very difficult to break it. (Student 8, high-improvement learner; comment on a run-on sentence, corpus group)</td>
</tr>
<tr>
<td>Not understanding the feedback at all</td>
<td>Unsuccessful error correction in the post-test due to not learning about error noticing and correction with the feedback</td>
<td>I was not learning because “cl” is not clear. I don't know why it is a “cl” error where it is, and how to correct it. So I can't correct it in the paper [in the post-test]. (Student 11, low-improvement learner; comment on a subordinate clause error, code group)</td>
</tr>
<tr>
<td>Incomplete understanding of the feedback about error correction</td>
<td>Unsuccessful error correction in the post-test due to incomplete understanding of the feedback</td>
<td>I know the sentence sounds weird, but I am not sure where exactly the error is... [Researcher: Do you think you learned anything from the feedback?] I think so, so I can find it, but I'm not sure how to correct it. (Student 13, intermediate-improvement learner; comment on a subordinate clause error, code group)</td>
</tr>
</tbody>
</table>

The primary and second researchers used the information in Table 3.10 to code idea units in four transcripts, and Cohen Kappa was .842. Transcripts remaining from the interview were coded by the primary researcher. Those codes were then used to construct seven themes on the reasons for successful and unsuccessful error correction in the post-
test: (1) learning from the feedback (learning from the feedback), (2) having already known how to correct target errors (pre-existing knowledge), (3) practicing personal writing (personal practice), (4) failure in error noticing (not noticing errors without feedback or hints), (5) forgetting what had been learned through feedback (forgetting), (6) the complexity of sentence structures in the post-test (complex sentences and long sentences), and (7) not learning with feedback or incomprehension of the feedback (not understanding the feedback at all and incomplete understanding of feedback on error correction). To compare students’ perceptions of knowledge learning with the corpus-based and coded feedback, the emergent themes and corresponding examples of idea units from the interview transcripts were analyzed and discussed. The percentage and number of idea units used to construct each theme and subtheme were reported as well. Qualitative data from the semi-structured interviews about students’ perceptions of learning about each type of error with the target feedback were also analyzed. From the two groups, students’ votes on the most and least difficult errors for error correction in the post-test were counted and compared, and the reasons behind those evaluations were extracted for discussion.

Research Question 4: What evidence suggests that corpus-based feedback is more appropriate than coded feedback for enabling error correction in advanced-low ESL learners?

Both quantitative data, collected with nine Likert-scale questionnaire items, and qualitative data, collected from semi-structured interviews, were analyzed to explore the learner fit quality of target corrective feedback for language learners. Among the nine items on learning impact in the questionnaire, two items that were negatively worded were
reversed scored. Cronbach’s alpha reliability of nine Likert-scale questions on learner impact was reported in Chapter four. Additionally, Cronbach’s alpha reliabilities of three questions for each sub-construct, namely clarity of feedback, usefulness of feedback, and self-correction with feedback, were also reported. Two groups of students’ answers to nine questionnaire items and the total scores of each construct were presented using descriptive statistics. To explore the difference between the perceptions of students receiving the two types of corrective feedback on each sub-construct of learning impact, a one-way ANOVA test was used for data analysis.

To provide in-depth analysis, qualitative data about the learner fit of the feedback from the semi-structured interview were analyzed for possible themes, subthemes, and examples. At the open coding stage, the primary and second researchers individually coded two interview transcripts on learner fit. Then, they discussed the codes and reached a consensus on code use for qualitative data analysis. Each code, coupled with its explanations and examples, is presented in Table 3.11.

Table 3.11. Code use, code description, and examples related to learner fit with target feedback.

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensibility of the feedback</td>
<td>Learners’ perceptions of whether the feedback is understandable to them</td>
<td>1. I can understand it pretty well. The words used in the examples are easy and the explanations are clearly. (Student 7, high-improvement learner, corpus group)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. I don’t understand the feedback. They are very vague to me. I think more explanations were needed. (Student 15, intermediate-improvement learner, code group)</td>
</tr>
</tbody>
</table>
Table 3.11. continued

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriateness of the feedback for error</td>
<td>Learners’ perceptions of the helpfulness of the feedback for completing error</td>
<td>The feedback is appropriate. I feel that those errors have been taught in the high school or</td>
</tr>
<tr>
<td>correction exercises</td>
<td>correction exercises</td>
<td>when I prepared exams, so the feedback reminds me. (Student 18, high-improvement learner,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>code group)</td>
</tr>
<tr>
<td>Individual needs</td>
<td>Learners’ perceptions of the feedback on meeting individual needs for learning</td>
<td>1. I feel that I like the feedback meets my needs. I like listening, but I don’t usually</td>
</tr>
<tr>
<td></td>
<td></td>
<td>get spoken feedback, so this time, I have choices. (Student 6, intermediate-improvement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>learner, corpus group)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Not good for me. It may be good for people who knows those errors well, but not for me.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I wanted the feedback supports me to learn, not leave me there wondering what’s missing the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sentence or how to correct it. (Student 15, intermediate-improvement learner, code group)</td>
</tr>
<tr>
<td>Individual learning pace</td>
<td>Learners’ perceptions of whether the feedback facilitates students’ learning</td>
<td>Yes. I don’t learn quickly. I need to read and think, and sometimes to read again. The</td>
</tr>
<tr>
<td></td>
<td>at their own pace</td>
<td>feedback helped me do my style. (Student 4, intermediate-improvement learner, corpus group)</td>
</tr>
</tbody>
</table>

The primary and second researchers used the information in Table 3.11 to code idea units in the four transcripts. Cohen’s Kappa was .880. The codes were then used to develop three themes on learner fit: clarity of feedback (understanding of the feedback), appropriateness of feedback for task completion (appropriateness of the feedback for error
correction exercises), and suitability for individual learner characteristics (individual needs and individual learning pace). To compare students’ perceptions of learner fit about the corpus-based and coded feedback, the emergent themes and corresponding examples of idea units from the interview transcripts were analyzed and discussed. The primary researcher then coded the remaining transcripts from interviews and categorized those codes into three themes on learner fit. The percentage and number of idea units used to construct each theme were reported as well.

Research Question 5: What evidence suggests that students treated with corpus-based feedback have a more positive learning experience than those treated with coded feedback?

The impact of learning about syntactic errors via coded and corpus-based feedback was explored with both quantitative and qualitative methods. Quantitative data were collected on 12 items from the Likert-scale questionnaire and qualitative data were collected from semi-structured interviews. Among the 12 items on learning impact in the questionnaire, two items that were negatively worded were reversed scored. Cronbach’s alpha reliability of the 12 Likert-scale questions on learner impact was reported. Additionally, Cronbach’s alpha reliability of three questions for each sub-construct, namely happiness about using feedback on other errors, satisfaction of learning experience, error awareness, and motivation for error correction, were also reported. Two groups of students’ answers to the 12 questionnaire items were presented using descriptive statistics. To explore the difference between the perceptions of students receiving the two types of corrective feedback on each sub-construct of learning impact, a one-way ANOVA test was conducted for data analysis.
Qualitative data from the semi-structured interviews were analyzed to explore the impact of learning. The primary and second researchers first worked separately on the open coding for the two interview transcripts on learning impact. When this first step was completed, the researchers then discussed the codes before reaching a consensus on code use in the data analysis, as shown in Table 3.12.

Table 3.12. Code use, code description, and examples for each code on learning impact quality

<table>
<thead>
<tr>
<th>Code use</th>
<th>Code description</th>
<th>Examples from the transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction of learning experience</td>
<td>Students’ expectations having been fulfilled by the learning experience with target feedback</td>
<td>Yes, I enjoy learning with the feedback. (Student 4, intermediate-improvement learner, corpus group)</td>
</tr>
<tr>
<td>Enthusiasm for using target feedback to correct other errors</td>
<td>Students’ enthusiasm at the prospect of using target feedback to correct other types of errors in academic writing</td>
<td>Yes, I would like [receive the feedback on other errors]. If I can receive feedback [on other errors] like this in the future, I will learn more knowledge. (Student 6, intermediate-improvement learner, corpus group)</td>
</tr>
<tr>
<td>Motivation for error correction by oneself</td>
<td>Students’ intrinsic motivation to correct the three target types of syntactic errors in the future</td>
<td>I paid more attention to those [syntactic errors] now. I don’t know how do you design the exercises, but it includes many mistakes I often make. So when I work on my draft, I will not submit it after finishing the first draft. I feel that I should correct those errors first. (Student 14, intermediate-improvement learner, code group)</td>
</tr>
<tr>
<td>Motivation for error correction with teacher feedback</td>
<td>Students’ intrinsic motivation to correct errors when receiving teacher feedback</td>
<td>When teacher highlighted my sentence as a fragment, I want to make sure I correct them correctly. [Researcher: Is what you are doing now different from your previous revision?] I previously don’t take it seriously. I just go and correct it. Now, because I learned, I want to make sure it’s correct. (Student 1, low-improvement learner, corpus group)</td>
</tr>
<tr>
<td>Code use</td>
<td>Code description</td>
<td>Examples from the transcripts</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Awareness of errors in one’s own writing</td>
<td>The improvement of students’ ability to notice the three target types of syntactic errors in their own academic writing</td>
<td>I think so. I was unaware of my mistakes. For example, I used to use “against” as a verb, but now I know it’s not. I should say “be against something.” (Student 13, intermediate-improvement learner, code group)</td>
</tr>
<tr>
<td>Awareness of errors in other people’s writing</td>
<td>The improvement of students’ ability to notice the three target types of syntactic errors in other people’s academic writing</td>
<td>I found I can notice more errors. Like last time, I read my classmate’s paper, I told him “you have many run-on sentences in your paper” and explained to him where they are. (Student 15, intermediate-improvement learner, code group)</td>
</tr>
<tr>
<td>Awareness of the importance of correcting errors in academic and formal writing</td>
<td>The improvement of students’ awareness of the importance of correcting the three target types of errors in academic and formal writing</td>
<td>I used to write as what I say. Now I know it’s very important not to write run-on in my paper. I need to separate them or use words to make connections. (Student 4, intermediate-improvement learner, corpus group)</td>
</tr>
</tbody>
</table>

The primary and second researchers used the information to code idea units in the four transcripts. Cohen’s Kappa was .923. From here, we grouped the code into three main themes: affective impact (satisfaction of learning experience and enthusiasm for using the target feedback on other errors), intrinsic impact (motivation for error correction by oneself and motivation for error correction with teacher feedback), and cognitive impact (awareness of errors in one’s own writing, awareness of errors in other people’s writing, and awareness of the importance of correcting errors in academic and formal writing). To gain insight into students’ perceptions of the impact of either type of corrective feedback, the relevant examples under each theme and their subthemes were discussed. The
percentages and number of idea units used to construct each theme and subtheme were reported as well.

3.6. Chapter Summary

This chapter provided detailed descriptions regarding the methodology employed in the dissertation. It first outlined the mixed methods using a convergent design to elucidate how this method facilitated understanding of the research problems. Information about the participants, the sampling procedure, and the teaching context illustrated the specific teaching and learning contexts used in the dissertation. Then, the presentation of data collection materials (i.e. error correction practice exercises, pre- and post-tests, the error awareness sheet, participants’ background survey, the Likert-scale questionnaire, instructions for the think-aloud protocols, and semi-structured interview questions) and data collection tools (the My Feedback website), coupled with an account of the data collection procedure, explained how the research was implemented and how the materials and tools were developed and used for data collection. Finally, the data coding and analyses for each research question were described, and the interrater reiabilities for quantitative and qualitative data analyses were reported. Based on the explanations on data analyses, the next chapter presents the findings for each of the three research questions about the language learning potential presented by the two types of corrective feedback.
CHAPTER 4. RESULTS ON THE LANGUAGE LEARNING POTENTIAL OF CALL CORRECTIVE FEEDBACK

In accordance with the logical framework for the development and evaluation of CALL corrective feedback, the first criterion investigated was language learning potential. This criterion was operationalized to examine three aspects: 1) error noticing using CALL corrective feedback, 2) error correction using CALL corrective feedback, and 3) knowledge gained from CALL corrective feedback. This chapter presents results derived from data concerning these three aspects of the language learning potential of corpus-based and coded feedback. The results are based on analysis of both quantitative and qualitative data. The quantitative data consist of the test scores of 90 participants (45 in the code group and 45 in the corpus group) on two error correction practice exercises, along with their pre- and post-test scores. The qualitative data are the think-aloud utterances and semi-structured interview answers of 18 participants selected from the original 90. Each of the following sections explores one aspect of the language learning potential, starting with the quantitative data analysis findings followed by the qualitative results, and the triangulation of both results for the purpose of interpretation.

4.1. Effects of Corpus-based and Coded Feedback on Error Noticing (Q1)

To examine the first research question, “What evidence suggests that students with corpus-based feedback are better equipped to notice syntactic errors than students with coded feedback?”, the quantitative data from 90 students’ error noticing scores on the two error correction exercises (each of which included 21 syntactic errors) and qualitative data from 18 students’ think-aloud protocols and semi-structured interviews were analyzed.
Overall, these data indicated that the corpus-based feedback was more effective than the coded feedback for facilitating error noticing. Both types of results were triangulated to develop a context-specific argument for the language learning potential of corpus-based and coded feedback on learning about syntactic errors.

4.1.1. Quantitative Data Results

The descriptive data regarding students’ scores for noticing errors on Error Correction Practice Exercises One and Two are included in Table 4.1.

Table 4.1 Descriptive statistics of participants’ total noticing scores on exercises one & two (N<sub>corpus</sub> = 45; N<sub>code</sub> = 45)

<table>
<thead>
<tr>
<th>Groups</th>
<th>No. of Participants</th>
<th>Exercise One</th>
<th>Exercise Two</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (out of 21)</td>
<td>SD</td>
</tr>
<tr>
<td>Corpus</td>
<td>45</td>
<td>18.33</td>
<td>1.88</td>
</tr>
<tr>
<td>Code</td>
<td>45</td>
<td>12.16</td>
<td>4.20</td>
</tr>
</tbody>
</table>

The total scores for noticing errors in the first exercise were higher in the corpus group (n=45, Mean= 18.33, SD= 1.88) than the code group (n=45, Mean= 12.16, SD= 4.20). On Exercise Two, the noticing scores of the corpus group (n=45, Mean= 18.87, SD= 1.43) were higher again than those in the code group (n=45, Mean= 11.71, SD= 4.77).

Given the dependent variable was a continues variable (i.e., noticing scores), measured from 0-21, and the independent variable was a nominal variable (i.e., feedback type), which has two groups: corpus group and code group, a one-way ANOVA is appropriate to determine whether there were any statistically significant differences between the means of two independent groups. As assessed by inspection of the boxplot in Figure 4.1, there were no outliers in the data. Therefore, the assumptions for ANOVA were tested.
Figure 4.1 *Boxplot of error noticing scores on error correction exercise one (Total Scores= 21)*

The assumptions of normality and homogeneity of variance were tested. The scores of the corpus groups were not normally distributed ($p = .018 < .05$), while those of the code group were normally distributed ($p = .721 > .05$), as assessed by Shapiro-Wilk's test of normality. According to Field (2013), if the sample size of each group is larger than 30, the ANOVA is robust when the data are not normally distributed. The assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p = .000 < .05$). Welch's ANOVA was reported since this statistic is appropriate when the homogeneity of variance assumption is not met (Field, 2013). The score differences between the two groups were statistically significant (Welch's $F (1, 60.86) = 80.96, p = .000 < .0005$, partial $\eta^2 = .47$). The effect size indicated that 47% of variance in the error noticing scores can be explained by the treatment of two types of feedback in error correction exercise one. The result revealed that the scores of the corpus group (Mean= 18.33, SD= 1.88) were significantly higher than the scores of the code group (Mean= 12.16, SD= 4.20) in Exercise One.
A second ANOVA was conducted to assess students’ error noticing performance on the second exercise. There were no outliers in the data, as seen in Figure 4.2.

![Boxplot of error noticing scores on error correction exercise two (Total Scores= 21)](image)

Figure 4.2 Boxplot of error noticing scores on error correction exercise two (Total Scores= 21)

The scores of the corpus group (p = .114 > .05) and the code group (p = .189 > .05) were normally distributed when assessed with Shapiro-Wilk’s test of normality. However, when assessed using Levene’s test for equality of variances, the assumption of homogeneity of variance was violated (p = .000 < .05). The Welch’s F is appropriate to be reported when the homogeneity of variance is not met (Brown & Forsthe, 1974; Field, 2013). The one-way ANOVA with Welch test showed that the scores between the corpus and the code group were significantly different (Welch’s F (1, 51.88) = 93.01, p < .0005, partial $\eta^2 = .51$). The effect size indicated that 51% of variance in the error noticing scores can be explained by the treatment of two types of feedback in error correction exercise two.

In addition to the total scores, the noticing scores of each group on sentence fragments, run-on sentences, and subordinate clause errors are reported in Table 7 and visualized in chart 4.2.
Table 4.2 Descriptive statistics of participants’ scores for noticing three types of syntactic errors in exercises one & two (N<sub>corpus</sub> = 45; N<sub>code</sub> = 45)

<table>
<thead>
<tr>
<th>Errors</th>
<th>Groups</th>
<th>Mean (out of 7)</th>
<th>SD</th>
<th>Mean (out of 7)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRAG</td>
<td>Corpus</td>
<td>6.53</td>
<td>.69</td>
<td>6.69</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>3.93</td>
<td>1.85</td>
<td>4.07</td>
<td>2.02</td>
</tr>
<tr>
<td>RO</td>
<td>Corpus</td>
<td>5.84</td>
<td>1.14</td>
<td>6.04</td>
<td>.71</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>4.44</td>
<td>1.51</td>
<td>4.44</td>
<td>1.71</td>
</tr>
<tr>
<td>CL</td>
<td>Corpus</td>
<td>5.96</td>
<td>.82</td>
<td>6.13</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>3.78</td>
<td>1.73</td>
<td>3.20</td>
<td>1.91</td>
</tr>
</tbody>
</table>

FRAG: sentence fragment; RO: run-on sentence; CL: subordinate clause errors

Figure 4.3 Line graphs of mean scores for noticing three types of syntactic errors in Exercises One & Two (N<sub>corpus</sub> = 45; N<sub>code</sub> = 45)

It is noted that the corpus group outperformed the code group in all three error categories on the two error correction practice exercises. Students in the corpus group most successfully noticed sentence fragments, while students in the code group most successfully noticed run-on sentences than the other two types of errors.

4.1.2. Qualitative Data Results

Students’ utterances about their performance on noticing errors with target feedback in the think-aloud protocol and semi-structured interview were coded as idea units and are categorized in Table 8. Their comments were interpreted as positive when
they suggested that the target feedback was helpful for their error noticing and negative when they indicated that they failed to notice errors with the feedback. As seen in table 4.3, 96.48% of the idea units from the corpus group reflect a positive response to corpus-based feedback on error noticing; by contrast, 58.43% of the idea units from the code group represent a positive response to coded feedback.

Table 4.3 Idea units reflecting positive and negative evaluations of both types of feedback on error noticing from both groups (N_{corpus interviewee} = 9; N_{code interviewee} = 9)

<table>
<thead>
<tr>
<th>Feature Description</th>
<th>Corpus group (85 idea units)</th>
<th>Code group (89 idea units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive 96.48% (82 units)</td>
<td>Negative 3.52% (3 units)</td>
</tr>
<tr>
<td>- Enhanced noticing with colors, examples, and oral explanations</td>
<td>92.94% (79 units)</td>
<td>Needlessness of oral feedback 3.52% (3 units)</td>
</tr>
<tr>
<td>- Noticing confirmation</td>
<td>3.52% (3 units)</td>
<td>Enhanced noticing with codes 58.43% (52 units)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insufficient information for error noticing 37.08% (33 units)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not confirming error noticing 4.49% (4 units)</td>
</tr>
</tbody>
</table>

4.1.2.1 Students’ perception of the corpus-based feedback on error noticing

Based on the data analysis, 96.48% percent of utterances from the corpus group were positive about the effects of corpus-based feedback on error noticing, and 3.52% were negative.

Enhanced noticing with colors, examples and oral explanations. Students’ reflections on their ability to notice errors primarily consisted of references to three main features in the feedback: color use, correct and incorrect examples, and oral explanations. In corpus-based feedback, colors were used as notifications: red indicated the erroneous
parts of sentences, while green indicated their corresponding correction. All interviewees addressed the importance of color-coded feedback for their ability to notice errors.

a. Yeah, it’s like, with the feedback. Actually I really didn’t notice without the feedback and colors. (Student 3, low-improvement learner; comment on a subordinate clause error, corpus group)

b. Yeah, they tell me how to find the errors with colors. (Student 4, intermediate-improvement learner; comment on a sentence fragment, corpus group)

c. The red colors shows me the sentence is wrong and where is wrong. (Student 7, high-improvement learner; comment on a run-on sentence, corpus group)

d. It’s easy. While I looked at red colors, I know the problem. (Student 9, high-improvement learner; comment on a run-on sentence, corpus group)

Students also positively commented on the examples used in the feedback for error noticing.

a. I compared the incorrect to the correct ones. Then I compared them to my sentence [to notice]. (Student 6, intermediate-improvement learner; comment on a run-on sentence, corpus group)

b. Examples are helpful [for noticing]. With it, I can see the problem. (Student 1, low-improvement learner; comment on a run-on sentence, corpus group)

c. Yes, it does [help me notice the errors]. Actually, the example does. (Student 9, high-improvement learner; comment on a subordinate clause error, corpus group)

d. The examples tell me. Like this one, I know the word “by” is not good here. (Student 8, high-improvement learner; comment on a sentence fragment, corpus group)

Two students reported a heightened ability to notice errors as a result of listening to oral explanations provided in the feedback:

a. I listened and read at the same time. [Do you think the oral feedback is helpful for you to notice the error?] Definitely. I feel listening is more helpful. The incorrect part was emphasized, so I can easy notice and correct. (Student 4, intermediate-improvement learner; comment on a run-on sentence, corpus group)
b. I’m an aural learner, so listening to the feedback is more efficient to understand where the error is and how to correct it. (Student 6, intermediate-improvement learner; comment on a subordinate clause error, corpus group)

Noticing confirmation. One student observed that the feedback helped confirm her own identification of errors, a feature that she found useful:

I looked at the sentences, and search for errors. And I know “there are” is wrong and should be deleted. I looked at the feedback it tells me the same thing. (Student 5, intermediate-improvement learner; comment on a subordinate clause error, corpus group)

In addition to the positive comments on written and oral corpus-based feedback, students also made negative or neutral comments about the oral explanations.

The needlessness of oral explanation. Three students felt that the written feedback alone satisfied their needs for noticing errors and that the accompanying oral explanation was unnecessary.

a. [Researcher: Do you think the spoken feedback is useful for you to notice errors?]
   Not really. I feel I can see errors when reading the feedback. I listened once, and I don’t think I need to listen. (Student 8, high-improvement learner; comment on errors in general, corpus group)

b. I usually do not listen cause it’s unnecessary, colors and [exemplified] sentences are good. (Student 7, high-improvement learner; comment on errors in general, corpus group)

4.1.2.2. Students’ perception of the coded feedback on error noticing

The comments on error noticing provided by students in the code group were 58.43% positive and 41.57% negative.

Enhanced noticing with codes. Students valued the codes in the feedback because they were able to use them as hints to locate errors.

a. Yes, it offers me some hints for finding errors. It’s always helpful. (Student 17, high-improvement learner; comment on a sentence fragment, code group)
b. Yes. I can [find the error with the coded feedback]. It says “RO,” so I know the whole sentence includes different clauses [that] need to be separated. (Student 12, low-improvement learner; comment on a run-on sentence, code group)

c. The feedback says it’s a “FRAG.” I found the sentence doesn’t have a verb, so should add one. (Student 16, high-improvement learner; comment on a sentence fragment, code group)

d. Yeah, this one, it also helped. The feedback tells what type of errors in the sentence. I read it again, like this one, sometimes one more time to find it. (Student 14, intermediate-improvement learner; comment on a run-on sentence, code group)

e. Yeah, because the sentence sounds good to me. I won’t catch it without the feedback. (Student 17, high-improvement learner; comment on a subordinate clause error, code group)

By comparison, 41.57% of the utterances from the code group reflected students’ negative view of the coded feedback in helping them notice syntactic errors. Students found the coded feedback included limited information that was insufficient to assist them in noticing target errors:

a. It’s not really helpful. It says “RO”, but it sounds correct and I can’t find errors. (Student 10, low-improvement learner; comment on a run-on sentence, code group)

b. The feedback is too general. I’m trying to find errors, but it didn’t give me clues. (Student 15, intermediate-improvement learner; comment on a subordinate clause error, code group)

c. No, it doesn’t give me any help. (Student 12, low-improvement learner; comment on a subordinate clause error, code group)

d. I need some examples and explanations. It’s too unclear and not helpful. (Student 13, intermediate-improvement learner; comment on a subordinate clause, code group)

Students also reported uncertainty as to whether they had correctly identified errors. The feedback failed to inform them if they had wrongly identified an error.

a. A little bit. I would like it to be more specific. Cause I found the errors, but if my thought is wrong, I changed it incorrect. Then it doesn’t help and I didn’t learn from the feedback. Probably I will do the same thing in my writing. (Student 17, high-improvement learner; comment on run-on sentence, code group)
b. It helped somehow. [Researcher: Why?] It tells that it’s a fragment sentence. I think a “period” is wrong here, but that’s my guess. I don’t know it’s correct or not. [Researcher: Is it important to know precisely where the error is?] I think so, because I may see something correct as incorrect and miss the incorrect part. (Student 15, intermediate-improvement learner; comment on a sentence fragment, code group)

After assessing the helpfulness of the target feedback for noticing syntactic errors in general, students also discussed the effects of both types of feedback for noticing sentence fragments, run-on sentences, and subordinate clause errors respectively. The chart below presents the number of students from each group who perceived the target feedback as helpful for noticing each error type.

![Chart showing the number of students who perceived the target feedback as helpful for noticing each error type](image)

Figure 4.3 *Number of students who perceived the target feedback as helpful for noticing each error type (N_{corpus-group interviewee} = 9; N_{code-group interviewee} = 9)*

4.1.2.3. Students’ perception of the corpus-based feedback on noticing sentence fragments, run-on sentences, and subordinate clause errors

As illustrated in Figure 4.4., all nine students in the corpus group perceived the feedback as being useful in helping them notice sentence fragments, and eight were satisfied with the feedback for noticing run-on sentences and clause errors. The reasons that they offered are very similar: the use of colors, examples, and oral explanations helped them notice erroneous parts of the sentences:
a. It’s helpful [to find sentence fragments]. The colors and examples, all of those are good. It’s very direct. [Researcher: What about run-on sentences?] The same, very good. [Researcher: And clause errors?] I think they are equally good. (Student 9, high-improvement learner, corpus group)

b. Feedback are good for all. So many times, each sentence needs to be corrected in more than one place. The comments [in the feedback] use “red” to mark all the errors in the sentences. Because of this, I can find all of them. (Student 3, low-improvement learner, corpus group)

c. I won’t ask more. The feedback help me notice errors [in sentence fragments]. Especially I’m listening to the feedback. [Researcher: What about run-on sentences?] I can. [Researcher: What about clause errors?] I can find it, too. (Student 4, intermediate-improvement learner, corpus group)

d. Many times, I can find many errors by myself. When I can’t, the feedback helps me find the rest. I think they are good feedback. (Student 8, high-improvement learner, corpus group)

Among those nine students, one had positive perceptions of the helpfulness of the corpus-based feedback for noticing sentence fragments, but had reservations about the feedback for noticing run-on sentences and subordinate clause errors:

a. The feedback helps me for most run-ons, but when the sentences are very very long, I sometimes cannot find all errors in the sentence. Like this, I think I wasn’t noticing the last part is still a “RO.” (Student 2, low-improvement learner, comment on a run-on sentence, corpus group)

b. The same thing for the clauses. They are complex and long. It [the feedback] only helps me find one part. (Student 2, low-improvement learner, comment on a subordinate clause error, corpus group)

4.1.2.4. Students’ perception of the coded feedback on noticing sentence fragments, run-on sentences, and subordinate clause errors

Students in the code group commented on the helpfulness of the coded feedback for noticing individual types of syntactic errors. Six out of nine students found that the feedback helped them notice run-on sentences, four found the feedback as being useful in
helping them notice sentence fragments, while three remarked on the helpfulness of the feedback in noticing subordinate clause errors.

**Sentence fragments.** Four out of nine students found that the error codes helped them locate errors in sentences.

a. The word “Frag” is good to notice. When I see it, I try to look for subject or verbs in the sentence. (Student 17, high-improvement learner, comment on sentence fragments, code group)

b. Sentence fragments is the easiest one. I can [find the error with the feedback]. (Student 14, intermediate-improvement learner, comment on sentence fragments, code group)

However, the remaining five students indicated that they had difficulties in noticing errors. They either said they failed to understand the feedback or were unable to find the erroneous part of the sentence with the coded feedback.

a. I found difficult…. because I don’t understand why some are fragment sentence. So I can’t find it. (Student 10, low-improvement learner, comment on sentence fragments, code group)

b. I feel like I want to know what’s missing, but the feedback doesn’t tell. I’m confused. (Student 15, intermediate-improvement learner, comment on sentence fragments, code group)

**Run-on sentences.** Five students found the coded feedback to be effective in helping them notice run-on sentences.

a. It’s easy. I think notice it. (Student 14, intermediate-improvement learner; comment on run-on sentences, code group)

b. It’s [the feedback] good. I can find errors. (Student 13, intermediate-improvement learner; comment on run-on sentences, code group)

c. The feedback is alright for run-on. All I need to do with run-on is to separate them. (Student 18, high-improvement learner; comment on run-on sentences, code group)
The other three had reservations about the feedback. They felt that the feedback failed to illustrate what made a sentence a run-on sentence, and that it was not sufficient to help notice errors in sentences containing many commas or clauses.

a. I liked it, like this one, I can find the “comma” should be changed to “period.” But for this one, I don’t think it’s a run-on sentence. (Student 16, high-improvement learner; comment on run-on sentences, code group)

b. Not very good. Something like put sentences into small sentences. If there’s too many commas, I don’t know should I stop here, or there, or there. Student 10, low-improvement learner; comment on run-on sentences, code group)

**Subordinate clause errors.** Fewer people perceived the coded feedback as being helpful for noticing subordinate clause errors: only three students praised the feedback for its usefulness in helping them detect clause errors.

a. I personally found it helpful. I made some clause errors in my writing, but the feedback is like a reminder and helper. It always highlight the sentence and I found it easily. (Student 18, high-improvement learner; comment on a subordinate clause error, code group)

b. [Researcher: What about the clause errors? Can you find errors with the feedback?] I think so. (Student 17, high-improvement learner; comment on a subordinate clause error, code group)

However, the remaining six students found the feedback unhelpful, citing confusion over the codes or insufficient information.

a. It doesn’t teach me anything. So this is a clause error, then where? I don’t know. Student 12, low-improvement learner; comment on a subordinate clause error, code group)

b. It’s not good. I tried to find the errors by myself, but it’s too hard for me [to notice]. (Student 13, intermediate-improvement learner; comment on a subordinate clause error, code group)

c. The feedback said “cl.” There are a lot of clauses in this sentence, and I don’t know why part is problematic. (Student 11, low-improvement learner; comment on a subordinate clause errors, code group)
4.1.3. Triangulation of Quantitative and Qualitative Results

Data for the investigation of the first aspect of language learning potential were drawn from the performance of 90 participants (45 from each group) in noticing errors in the two error correction practice exercises and from think-aloud protocols and semi-structured interviews conducted with 18 participants (9 participants from each group). The quantitative results from one-way ANOVA tests showed that the corpus group outperformed the code group in noticing syntactic errors on the two error correction practice exercises. The error noticing scores of the corpus group were statistically significantly higher than those of the code group. In addition, the descriptive statistics of error noticing scores for individual syntactic errors showed that the corpus group attained higher scores on sentence fragments, run-on sentences and subordinate clause errors respectively than the code group; the corpus group had the best performance in noticing sentence fragment errors, while the code group had the best performance in noticing run-on sentences.

Students' perception of the effectiveness of the two types of feedback on error noticing mirrored their performance on the error correction practice exercises: 96.48% of the comments from the corpus group reflected positively on the helpfulness of the corpus-based feedback for error noticing; by contrast, 58.43% of the utterances from the code group were positive. Specifically, the positive comments from the corpus group remarked on various features of the feedback, such as the use of color, examples, and oral explanations, which students found to contribute to their success in noticing errors. By comparison, positive comments from the code group linked students’ successful performance to the error codes in the feedback; at the same time, 41.57% of the code
group’s comments reflected dissatisfaction with the coded feedback. Specifically, students wanted coded feedback to provide more detailed information, such as examples or explanations, to help locate errors and confirm their own conjectures.

When discussing each type of syntactic error in the interviews, eight out of the nine participants from the corpus group found corpus-based feedback to be sufficient in helping them notice all three types of errors in the articles. One student remarked that, when dealing with long and complex sentences, he was unable to notice all the errors, despite using corpus-based feedback. The interviewees from the code group were not as satisfied as those from the corpus group. Six out of the nine participants from the code group were satisfied with the coded feedback for alerting them to run-on sentences; they felt run-on sentences only required them to focus on the interconnections among sentences, which made the errors relatively easily located. By comparison, four participants found the feedback helpful for noticing sentence fragments, and three found it helpful for noticing subordinate clause errors. The rest of the interviewees in the code group remarked that sentence fragments or subordinate clauses were complex: errors may occur in different places within sentences and were difficult to identify. This being the case, they found the feedback to be less detailed and comprehensible that they would have hoped. The perceptive insights complemented the quantitative data to provide an understanding of how detection occurred or failed to occur with each type of error and why corpus-based feedback was more helpful than coded feedback in helping students notice syntactic errors.

Students’ performance on the error correction practice exercises and their utterances in the think-aloud protocol and semi-structured interview also illuminated the quality of both types of feedback for error correction, as discussed in the following section.
4.2. The Effects of Coded and Corpus-based Feedback on Error Correction (Q2)

To answer the second research question, “What evidence suggests that students with corpus-based feedback are better equipped to correct syntactic errors than students with coded feedback?”, the quantitative data from 90 students’ error correction scores on the two error correction exercises and qualitative data from 18 students’ think-aloud protocols and semi-structured interviews transcripts were collected and analyzed. Overall, corpus-based feedback was more helpful for error correction than coded feedback from students’ performance and perception.

4.2.1. Quantitative Data Results

Regarding the quantitative measure, the descriptive data of students’ scores on Error Correction Practice Exercises One and Two are reported in Table 4.4.

Table 4.4 Descriptive statistics of participants’ total error correction scores on exercises one & two (N_{corpus} = 45; N_{code} = 45)

<table>
<thead>
<tr>
<th>Groups</th>
<th>No. of Participants</th>
<th>Exercise One</th>
<th>Exercise Two</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (out of 21)</td>
<td>SD</td>
</tr>
<tr>
<td>Corpus</td>
<td>45</td>
<td>17.50</td>
<td>1.91</td>
</tr>
<tr>
<td>Code</td>
<td>45</td>
<td>11.53</td>
<td>4.30</td>
</tr>
</tbody>
</table>

The total scores on the first exercise were higher in the corpus group (n=45, Mean=17.50, SD=1.91) than the code group (n=45, Mean=11.53, SD=4.30). Similarly, on Exercise Two, the total scores in the corpus group (n=45, Mean=18.49, SD=1.54) were higher than those in the code group (n=45, Mean=11.07, SD=4.70). A one-way ANOVA was conducted to evaluate whether students’ scores in the two groups were statistically different in Exercise One. As assessed by inspection of the boxplot in Figure 4.5., there were no outliers in the data.
Figure 4.4 Boxplot of error correcting scores on error correction exercise one (Total Scores=21)

The scores of the corpus group (p = .060 > .05) and the code group (p = .862 > .05) were normally distributed, as assessed by Shapiro-Wilk’s test of normality. The assumption of homogeneity of variances was violated, as assessed by Levene’s test for equality of variances (p = .000 < .05). Given that the assumption of homogeneity of variances was violated, Welch’s F test was used to report the results. The results with one-way ANOVA have shown that the score differences between the two groups were statistically significant (Welch’s F (1, 60.84) = 89.48, p < .0005, partial η² = .45). The effect size indicated that 45% of variance in the error correction scores can be explained by the treatment of two types of feedback in error correction exercise one. Accordingly, the results of Welch’s F revealed that the scores of the corpus group (Mean=17.50, SD=1.91) were significantly higher than the scores of the code group (Mean=11.28, SD=4.28) on Exercise One.
A second ANOVA was conducted to assess students’ performance on the second
exercise. There were no outliers in the data, as seen in Figure 4.6.

![Boxplot of error correcting scores on error correction exercise two](image)

**Figure 4.5 Boxplot of error correcting scores on error correction exercise two** ($N_{\text{corpus}} = 45; N_{\text{code}} = 45$)

The scores of the corpus group ($p = .141 > .05$) and the code group ($p = .164 > .05$) were normally distributed when assessed with Shapiro-Wilk’s test of normality. Moreover, when assessed using Levene’s test for equality of variances, the assumption of homogeneity of variance was violated ($p = .000 < .05$). Given that the assumption of homogeneity of variances was violated, Welch’s F test was used to report the results. The one-way ANOVA showed that the scores between the corpus and the code group were significantly different (Welch’s F (1, 53.339) = 100.949, $p < .0005$, partial $\eta^2 = .44$). The effect size indicated that 44% of variance in the error correction scores can be explained by the treatment of two types of feedback in error correction exercise two.

Students in the corpus group had greater success correcting sentence fragments, run-ons, and subordinate clause errors than those in the code group, as shown in Table 4.5. Figure 4.7, which provides a visual representation of the mean scores for correcting the
three types of syntactic errors, demonstrates that students from the corpus group used corpus-based feedback most successfully to correct sentence fragments, while students from the code group used coded feedback most successfully to correct run-on sentences. Both the data in the table and the plot lines in the chart show that the corpus group outperformed the code group when correcting all three types of errors.

Table 4.5 *Descriptive statistics of participants’ scores correcting all three types of syntactic errors in exercises one & two (N\text{\textsubscript{corpus}} = 45; N\text{\textsubscript{code}} = 45)*

<table>
<thead>
<tr>
<th>Errors</th>
<th>Groups</th>
<th>Exercise One</th>
<th></th>
<th></th>
<th>Exercise Two</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (out of 7)</td>
<td>SD</td>
<td>Mean (out of 7)</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>FRAG</td>
<td>Corpus</td>
<td>6.07</td>
<td>.79</td>
<td>6.44</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>3.64</td>
<td>1.87</td>
<td>3.80</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td>RO</td>
<td>Corpus</td>
<td>5.77</td>
<td>1.16</td>
<td>5.93</td>
<td>.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>4.10</td>
<td>1.49</td>
<td>4.18</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>Corpus</td>
<td>5.67</td>
<td>.87</td>
<td>6.11</td>
<td>.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>3.79</td>
<td>1.74</td>
<td>3.09</td>
<td>1.84</td>
<td></td>
</tr>
</tbody>
</table>

FRAG: sentence fragment; RO: run-on sentence; CL: subordinate clause errors

Figure 4.6 *Mean scores for correcting all three types of syntactic errors in exercises one & two (N\text{\textsubscript{corpus}} = 45; N\text{\textsubscript{code}} = 45)*

To explore reasons for students’ successful and unsuccessful error correction when using coded and corpus-based feedback, an analysis of qualitative data from think-aloud protocols and semi-structured interviews is provided in the following section.
4.2.2. Qualitative Data Results

Students’ utterances about their performance on correcting errors with target feedback in the think-aloud protocol and semi-structured interview were coded as idea units and are categorized in Table 4.6.

Table 4.6 Idea Units reflecting positive and negative evaluations of both types of feedback of error correction from both groups (\(N_{\text{corpus-group interviewee}} = 9; N_{\text{code-group interviewee}} = 9\))

<table>
<thead>
<tr>
<th></th>
<th>Corpus group (90 idea units)</th>
<th>Code group (85 idea units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87.78% (79 units)</td>
<td>12.22% (11 units)</td>
<td>46.88% (39 units)</td>
</tr>
<tr>
<td>Enhanced learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and understanding of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>errors</td>
<td>71.11% (64 units)</td>
<td>5.56% (5 units)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thinking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>enhancement and knowledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reminding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>46.88% (39 units)</td>
</tr>
<tr>
<td>Correction confirmation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.11% (10 units)</td>
<td>Confusion about grammatical</td>
<td>Failure to confirm students'</td>
</tr>
<tr>
<td></td>
<td>terms</td>
<td>attempts at correction</td>
</tr>
<tr>
<td></td>
<td>3.33% (3 units)</td>
<td>12.94% (11 units)</td>
</tr>
<tr>
<td>Usefulness of oral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.56% (5 units)</td>
<td>Needlessness of oral</td>
<td></td>
</tr>
<tr>
<td></td>
<td>feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.33% (3 units)</td>
<td></td>
</tr>
</tbody>
</table>

As seen in the table, 87.78% of the idea units from the corpus group reflect a positive response to corpus-based feedback; by contrast, 46.88% of the idea units from the code group represent a positive response to coded feedback. The following section contains a discussion of students’ positive and negative evaluations of both types of corrective feedback on error correction, with representative comments.
4.2.2.1. Students’ perception of corpus-based feedback on error correction

According to the data analysis, 87.78% percent of utterances from the corpus group were positive about the role of corpus-based feedback in error correction, while 12.22% were negative. Students who reported positive experiences with corpus-based feedback articulated three main reasons behind their appraisal: enhanced learning and understanding of errors, correction confirmation, and the usefulness of oral feedback in error correction.

Enhanced learning and understanding of error correction. All students in the corpus group appreciated corpus-based feedback primarily for its enhancement of their learning and understanding of errors and strategies for correction. The following statements are representative of students’ responses to the corpus-based feedback:

a. It told me where is wrong and why. This’s a learning process and I can understand why it’s wrong. (Student 1, low-improvement learner; comment on a run-on sentence, corpus group)

b. There’s an example given in the feedback and it’s really helpful to get to know... like it’s relatable to the sentence given in the paragraph and it helps a lot... Also, there are given three examples of correction and I know how to correct it. (Student 4, intermediate-improvement learner; comment on a subordinate clause error, corpus group)

c. I clicked on the cross. The window shows and told me it’s a sentence fragment. I need to add a verb here, because there’s no verb in this sentence. Because I don’t know “against” isn’t a verb, my [original] correction is wrong. The window.... The feedback, teaches me. (Student 5, intermediate-improvement learner; comment on a sentence fragment, corpus group)

d. I read the correct and incorrect sentences [in the feedback], and compare them to see the difference. Then I know that there are too many sentences with comma. I go back to my sentence and correct it. [Researcher: Do you think the feedback helped you correct the error?] Definitely. It shows me why it’s wrong, so I can correct it by myself. (Student 8, high-improvement learner; comment on a run-on error)
**Correction confirmation.** In addition to understanding and learning enhancement, four students felt that corpus-based feedback helped them to confirm their corrections.

a. I know the problems [in the sentence], but it’s always good to confirm it again [with the feedback]. (Student 9, high-improvement learner; comment on a run-on sentence, corpus group)

b. I kinda know how to correct it, but not very sure. Then I read the feedback and it’s what I thought, which is good. (Student 3, low-improvement learner; comment on a sentence fragment, corpus group)

c. I firstly changing “get” to “getting” because there are two verbs in the sentence. I read the feedback, which told me the same thing. So I think my change is correct. [Researcher: Do you think the feedback was helpful?] Yes, and it’s important to know I’m correct and why. (Student 5, intermediate-improvement learner; comment on a subordinate clause, corpus group)

d. I don’t know I’m correct [about my correction]. I go and check the feedback and it told me I’m correct, so I keep my changes. (Student 1, low-improvement learner; comment on a subordinate clause, corpus group)

**The usefulness of oral feedback.** Three students indicated that the use of oral feedback facilitated their understanding of problems in the sentences. When experiencing difficulties comprehending the written corpus-based feedback, they found that the oral feedback enhanced their understanding.

a. I read the feedback and try to understand them. Some terms, like this, is hard. I’m not sure if I’m understanding, so after listening to the audio, I can confirm if what I’m understanding is correct. The audio feedback is helpful. (Student 3, low-improvement group learner; comment on a run-on sentence)

b. I am a little confused about the [written] feedback. I listened to the feedback. The word “by” is unnecessary [in the examples in the feedback]. I look back the example in the feedback. I understand it quickly and go to the sentence and delete “by” in the sentence. (Student 6, intermediate-improvement learner; comment on a sentence fragment, corpus group)

c. The audio feedback explained it needs connectors between sentences. I read the example again and know the ways of correction. (Student 4, intermediate-improvement learner; comment on a run-on sentence, corpus group)
Student responses were not exclusively positive: 12.22% of the idea units (11 units) reflected a negative user experience with the corpus-based feedback on error correction. Students voiced their need for more examples and explanations, discussed their difficulties understanding the formal grammar terms, and suggested the needlessness of oral feedback in the corpus-based feedback.

**Insufficient examples and explanations in the feedback.** Even though many students expressed their appreciation of the examples in the feedback, two high-improvement learners desired more examples for subordinate clause errors.

a. I need more examples and explanations. (Student 7, high-improvement learner; comment on subordinate clause errors, corpus group)

b. I would like to see more examples that I can relate to, so I can learn it better. (Student 8, high-improvement learner; comment on subordinate clause errors, corpus group)

**Unfamiliarity with grammatical terms.** In addition, two students found themselves confused by the grammatical terms used in the corpus-based feedback.

Inclusion of this terminology was not helpful for their revision.

a. I don’t understand this [the term adjective pronoun]. It would be helpful if I understood it... maybe add some explanations. (Student 2, low-improvement learner; comment on subordinate clause error, Corpus group)

b. I know how to correct it using the example [in the feedback], but I don’t know what’s the meaning of “adverbial dependent clause.” [Researcher: Do you think it is helpful for your correction?] No, it’s not helpful... probably it’s my problem... because I’m not good at grammar, but for others, maybe... yes. (Student 3, low-improvement learner; comment on a clause error, corpus group)

**Needlessness of the oral feedback.** Although some students expressed a preference for oral feedback, two felt that the oral feedback was unnecessary.

a. I didn’t listen [to the oral feedback]. I understand how to correct it already, so I don’t spend more time listening the spoken feedback. (Student 8, high-improvement learner; comment on a sentence fragment, corpus group)
b. Basically, the oral feedback repeated what’s in the written feedback. I feel it’s not useful. (Student 2, low-improvement learner; comment on errors in general, corpus group)

4.2.2.2. Students’ perception of coded feedback on error correction

The code group articulated 46.88% positive comments on coded feedback for error correction and 53.12 % negative. Students enjoyed the coded feedback because they found that it enhanced their thinking and reminded them of their pre-existing knowledge:

a. The feedback said it’s a run-on. If I didn’t read the comment, I might just skip it [the error]…. It sounds good to me. While I think it again, I can find the error. Here, I will add a “and.” (Student 17, high-improvement learner; comment on a run-on sentence, code group)

b. Frag shows it’s a fragment sentence. It should missing [miss] something. I search in the sentence and here, I need to add a verb. It should be “was” here. (Student 14, intermediate-improvement learner; comment on a sentence fragment, code group)

c. It [the feedback] is useful. It reminded me of the knowledge I learned in the high school. (Student 18, high-improvement learner; comment on a run-on sentence, code group)

Negative evaluation about the coded feedback focused on its lack of information when explaining errors and providing solutions and its failure to confirm students’ attempts at error correction:

a. I know it is wrong, but I don’t know where it is wrong. I think I need more information for this sentence... I want to learn more about sentence structure, more details about fragment. (Student 11, low-improvement learner; comment on a sentence fragment, code group)

b. I don’t think I make it correct, but I don’t know how to correct it. Should I add “and” here or here. I need more information. (Student 16, high-improvement learner; comment on a run-on sentence, code group)

c. The sentence is good to me. Probably the feedback can explain why. (Student 13, intermediate-improvement learner; comment on a subordinate clause, code group)
c. After correct the error, I am not sure my revision was correct. (Student 15, intermediate-improvement learner; comment on a sentence fragment, code group).

d. The feedback is helpful, but I’m not very confident my revision is correct. (Student 12, low-improvement learner; comment on a subordinate clause error, code group)

In short, the corpus group’s positive evaluations of the corpus-based feedback outstripped the code group’s positive evaluations of the coded feedback in general. This result was consistent with students’ performances on the two error correction exercises, in which students with corpus-based feedback were more successful in correcting syntactic errors than students with coded feedback. The reasons that participants gave for their response to the feedback convincingly illustrate why students found corpus-based feedback helpful for their error correction, and what difficulties students encountered when using the coded feedback.

In addition to the evaluation of the effects of target corrective feedback in the error correction exercises, students also discussed their perception of the helpfulness of feedback when correcting each type of syntactic error. The following chart (Figure 4.8.) displays the number of students in each group who perceived the target feedback as being helpful for correcting each error type.

Figure 4.7 Number of students perceiving the target feedback as helpful for correcting each error type ($N_{corpus\text{-}group\ interviewee} = 9; N_{code\text{-}group\ interviewee} = 9$)
4.2.2.3. Students’ perception of corpus-based feedback on correcting sentence fragments, run-on sentences, and subordinate clause errors

As seen in the chart, all nine students in the corpus group found the feedback helpful for sentence fragment correction, eight of them found it helpful for run-on sentences, and seven found it helpful for subordinate clause errors. Among those nine students, two made positive comments about the helpfulness of the feedback for all three types of syntactic errors. The following quotations from those two participants illustrate their attitudes.

a. I think a person having problems with a specific error, all the feedbacks are really perfect. It gives you the whole idea of what you’re making a mistake about. (Student 9, high-improvement learner, corpus group)

b. All of them are helpful. I learned with feedback on all those three errors. I especially like the feedback on sentence fragment because I usually make such mistakes in the exercise, but now I know why they’re wrong. And I also like the other two [feedback on run-on sentences and clause errors]. All feedback are detailed and touch the points. (Student 5, intermediate-improvement learner, corpus group)

The remaining seven students from the corpus group discussed their satisfaction or dissatisfaction about the feedback on individual errors.

Sentence fragments. The remaining seven participants felt that the corpus-based feedback taught them how to deal with sentence fragments successfully.

a. It’s helpful because it teaches me how to correct it directly. I understand why, so it’s good to have such feedback. (Student 1, low-improvement learner, comment on sentence fragments, corpus group)

b. Yes, it is. I always have difficulties correcting fragments. When the sentence is long, I can’t find what’s missing, but the feedback helps find and correct the errors. (Student 4, intermediate-improvement learner; comment on sentence fragments, corpus group)

Run-on sentences. Out of the remaining participants, six liked the feedback on correcting run-on sentences:
a. It’s easy to read and understand. (Student 3, low-improvement learner; comment on run-on sentences, corpus group)

b. [Researcher: What about the feedback for run-ons and clause errors?] I think they are all good. As I said, they explained to me clearly, really like a tutor in my life. Easy to follow and remember. (Student 8, high-improvement learner; comment on run-on sentences, corpus group)

One student mentioned a negative aspect of corpus-based feedback on run-on sentences.

Some expressions [terms] used to explain run-on sentences in the writing makes me confused. I’m not familiar with those expressions. And probably need more explanations. (Student 2, low-improvement learner; comment on run-on sentences, corpus group)

**Subordinate Clause errors.** Five students liked corpus-based feedback on correcting clause errors. They explained their responses thus:

a. I feel it [clause errors] are difficult because I may need to change more than one words in the sentences... and their order. The feedback explains very clearly and I can imitate the examples to correct it. (Student 1, low-improvement learner; comment on subordinate clause errors, corpus group)

b. The explanations [for clause errors] are thorough. It teaches me different ways of correction, which I like most. (Student 4, intermediate-improvement learner; comment on subordinate clause errors, corpus group)

Two students expressed a need for more examples in the feedback on clause errors.

a. The feedback overall is good, but clause errors were complicated. I hope to include more examples and corrections. (Student 7, high-improvement learner; comment on subordinate clause errors, corpus group)

b. I wish having more examples. I feel more examples are always beneficial. (Student 8, high-improvement learner; comment on subordinate clause errors, corpus group)

**4.2.2.4. Students’ perception of coded feedback on correcting sentence fragments, run-on sentences, and clause errors**

Nine students from the code group also discussed their opinions on coded feedback for correcting the three types of syntactic errors: five students in the corpus group found the feedback helpful for run-on sentences, four of them found it helpful for sentence
fragment correction, and two found it helpful for subordinate clause errors. One student expressed his satisfaction with the coded feedback on correcting all three syntactic errors in general:

I think it’s good for all. It's very straightforward and I can easily use it for correction. (Student 17, high-improvement learner, code group)

The remaining eight discussed their positive and negative perceptions of the helpfulness of the coded feedback on correcting each individual error.

**Sentence fragments.** Among the remaining eight people, three found the coded feedback on sentence fragments to be helpful in reminding them of their pre-existing knowledge and teaching them ways of correcting errors.

a. The feedback is helpful because it reminds me of what I’ve learnt. (Student 18, high-improvement learner, comment on sentence fragments, code group)

b. The feedback [on sentence fragments] is good because when I see it, I look for something missing in the sentence, like subject, verb. So I can correct it by myself. (Student 14, intermediate-improvement learner, comment on sentence fragments code group)

The other five students held different opinions. They had expected the coded feedback to be more explicit and informative.

a. This feedback can always help me know what has happened in the sentence, but if I want to correct the sentence, I have to figure out what I need to add. But this [is what] feedback cannot give me. So, for the correction, it's not that easy to help me. (Student 15, intermediate-improvement learner; comment on sentence fragments).

b. When a sentence is short, it is easy to find and correct [with the feedback]. If the sentences is long and complicated, it's hard for me. Probably more explanations and information are needed. (Student 12, low-improvement learner; comment on sentence fragments, code group)

**Run-on sentences.** Participants made more positive comments about the feedback for correcting run-on sentences than for sentence fragments: four people felt that they were comfortable correcting run-on sentences with the coded feedback. Informed by the
feedback, they were able to separate clauses into independent sentences or add appropriate connectors.

a. Run-on is easy to correct [with the feedback] cause you only need to focus on something between sentences [punctuation]. It’s common sense. (Student 13, intermediate-improvement learner; comment on run-on sentences, code group)

b. Yes, I can notice and correct errors with the feedback. I only to pay attention to “comma” between two sentences and changed them to period. (Student 14, intermediate-improvement learner; comment on run-on sentences, code group)

The remaining four, however, found the feedback insufficient for their corrections:

a. The feedback is good for base form [of run-on sentence], like two sentences with a comma. The error is very easy to see. It is not good for long sentence. For example, this sentence is long. I think [I would] add an “and” after the second comma. But I’m not sure if I should add anything after the first comma. (Student 16, high-improvement learner; comment on run-on sentences, code group)

b. Not helpful, I think. I tried to correct, but I am not sure it’s correct. I always want some examples. (Student 10, low-improvement learner; comment on run-on sentences, code group)

Subordinate Clause errors. One student found the feedback useful. She found that the feedback provided useful guidance on his correction. Seven students were either confused about the meaning of “clause error” or complained that the feedback was insufficient to help with their corrections.

a. First of all, I don’t know the actual meaning of the clause, and I think my grammar has some problems. (Student 11, low-improvement learner; comment on subordinate clause errors, code group)

b. I can’t tell where is wrong and how to correct it. (Student 13, intermediate-improvement learner; comment on subordinate clause errors, code group)

c. There are many types of clause errors. I can’t figure out the correct version with it. (Student 15, intermediate-improvement learner; comment on subordinate clause errors)
d. I don't like the feedback. I try to find errors in the sentence, but the feedback didn’t tell me. I just correct it anyway. (Student 16, high-improvement learner; comment on subordinate clause errors, code group)

4.2.3. Triangulation of Quantitative and Qualitative Results

The overall findings concerning students’ error correction showed corpus-based feedback to be more effective in helping students correct syntactic errors than coded feedback. Specifically, the quantitative data from 90 students’ performance on Error Correction Exercises One and Two revealed that the scores of the corpus group were significantly higher than those of the code group; moreover, the correction scores of the corpus group for each error type (i.e. sentence fragments, run-on sentences, and subordinate clause errors) were also higher than each of those for the code group. Students in the corpus group were most successful at correcting sentence fragments; students in the code group were most successful correcting run-on sentences.

The think-aloud protocol and semi-structured interviews closely examined students’ perceptions of the two types of corrective feedback and reasons for students’ successful and unsuccessful error corrections. The corpus group generated 90 comments on the feedback, with 87.78% positive and 12.22% negative; students found the corpus-based feedback effective because it enhanced their understanding of the feedback and confirmed their corrections. By comparison, the code group uttered 85 comments, with 46.88% positive and 53.12% negative. Students reported that the coded feedback helped enhance their thinking and reminded them of their pre-existing knowledge. At the same time, 76.09% of students’ utterances faulted the coded feedback for containing insufficient information, and for generating confusion. In terms of the effects of the target feedback for enabling different types of error corrections, seven interviewees from the corpus group
found the feedback helpful for sentence fragments, run-on sentences, and subordinate clauses. They remarked that the feedback included detailed explanations and understandable and imitable examples offering clear guidance in correcting all three types of syntactic errors. Still, two interviewees had reservations about the corpus-based feedback on subordinate clause errors: they expressed their desire for more explanations and examples about subordinate clauses. In the code group, five interviewees found the feedback helpful for correcting run-on sentences, four found it useful for correcting sentence fragments, and only one found it helpful for correcting subordinate clause errors. The rest of the interviewees in the code group reported that the feedback failed to provide sufficient information for them to correct those three types of errors. The qualitative data illustrate the reasons for the corpus-based group’s higher scores correcting syntactic errors generally and individual errors specifically. The following section concerns the helpfulness of the feedback on knowledge learning.

4.3 The Effects of Coded and Corpus-based Feedback on Knowledge Application (Q3)

The third research question is “What evidence suggests that students treated with corpus-based feedback are better equipped to apply syntactic knowledge to correct errors in a new text than students treated with coded feedback?” Evidence that learners transferred knowledge to the correction of errors in new texts was obtained from all 90 participants in pre- and post-tests and 18 interviewees’ answers to semi-structured interview.
4.3.1 Quantitative Data Analysis

The descriptive data for the pre- and post-tests are presented in Table 4.7. As shown, the scores in the code group increased from 3.87 to 6.73, and those in the corpus group improved from 3.82 to 11.33 from the pre-test to the post-test.

Table 4.7 Descriptive statistics of pre- and post-tests ($N_{corpus} = 45; N_{code} = 45$)

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>Mean (out of 21)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Corpus</td>
<td>3.82</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>3.87</td>
<td>3.58</td>
</tr>
<tr>
<td>Post-test</td>
<td>Corpus</td>
<td>11.33</td>
<td>4.30</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>6.73</td>
<td>4.78</td>
</tr>
</tbody>
</table>

To further analyze students’ performance on the pre- and post-tests, a two-way mixed analysis of variance (ANOVA) was conducted with one within-subject factor—time (pre- and post-test)—and one between-subject factor—group (the code group and the corpus group). As shown in Figure 4.9, there were three outliers among the pre-test scores from the code group, three outliers among the pre-test scores from the corpus group, and one outlier among the post-test scores from the corpus group, as assessed by inspection of a boxplot. The post-test scores of the corpus group were normally distributed ($p = .443 > .05$), while the pre-test scores of the code group ($p = .000 < .05$) and corpus group ($p = .000 < .05$) and the post-test scores of the code group ($p = .014 < .05$) were not, when assessed by Shapiro-Wilk’s test. The assumptions of homogeneity of variances ($p_{pre} = .744$ $P_{post} = .083$) and covariance ($p = .694$) were met, as assessed by Levene’s test of the homogeneity of variances and Box’s test, respectively. Since there were many outliers in the data and the assumption of normality was not met, the results with the original data with the outliers and without outliers were reported and compared.
With the original data, there was a statistically significant interaction between feedback treatment and test scores in the pre- and post-tests ($F(1, 88) = 46.101$, $p < .0005$ partial $\eta^2 = .344$). The non-parallel lines in the profile plot (Figure 4.10.) show a statistically significant interaction effect, indicating that the two feedback groups performed differently over time.

There were no statistically significant differences in the pre-test scores between the two groups ($F(1, 88) = .013$, $p = .995$, partial $\eta^2 = .000$). Time was a significant factor:
students in both the code group (\(F(1, 44) = 32.049, p = .000, \text{ partial } \eta^2 = .421\)) and the corpus group (\(F(1, 44) = 266.762, p = .000, \text{ partial } \eta^2 = .858\)) improved significantly from the pre-test to the post-test. The effect sizes of the code and corpus groups indicated that 42.1% and 85.8% of variance in pre- and post-test scores can be explained by the time respectively. It should be noted that students in the two groups did not improve equally: there was a statistically significant difference in error correction success between the two feedback groups in the post-test (\(F(1, 88) = 23.035, p = .000, \text{ partial } \eta^2 = .207\)). The effect size indicated 20.7% of variances in the post-test scores can be attributed to the treatment of two types of feedback. Analysis indicates that, although students from both groups had significant improvement on their syntactic knowledge from the pre-test to the post-test, the corpus group improved more than the code group.

After all seven outliers were deleted from the original data, two new outliers from the pre-test in the code group emerged in the data set. Once these two new outliers were deleted from the data set, there were no outliers, as assessed by inspection of a boxplot. The data set included pre- and post-test scores from 41 participants from the corpus group and 40 participants from the code group. The data in the pre-test from the code group (\(p_{\text{code}} = .064\)) and in the post-test from the corpus group (\(p_{\text{corpus}} = .383\)) were normally distributed, but the pre-test scores of the corpus group (\(p_{\text{corpus}} = .001\)) and the post-test scores of the code group (\(p_{\text{code}} = .037\)) were not. The assumptions of the homogeneity of variances (\(p_{\text{corpus}} = .608; p_{\text{code}} = .053\)) and covariance (\(p = .069\)) were met. When the mixed two-way ANOVA was conducted with the new data, the statistically significant interaction between feedback treatment and test scores in the pre- and post-tests (\(F(1, 79) = 52.238, p < .0005 \text{ partial } \eta^2 = .398\)) were visible. No significant difference in the pre-test
scores between the two groups was observed ($F(1, 79) = .330, p = .567, \text{partial } \eta^2 = .004$).

Time was a significant factor: students in both groups—the corpus group ($F(1, 40) = 345.263, p = .000, \text{partial } \eta^2 = .896$) and the code group ($F(1, 39) = 33.733, p = .000, \text{partial } \eta^2 = .464$)—improved significantly from pre- to post-test. The effect sizes of the corpus and code groups indicated that 89.6% and 46.4% of variance in pre- and post-test scores can be explained by the time respectively. What is more, there was a statistically significant difference in successful error correction between the two feedback groups on the post-test ($F(1, 79) = 34.037, p = .000, \text{partial } \eta^2 = .301$). In summary, the data point to the same conclusion with or without the outliers. The effect size indicated 30.1% of variance in the post-test scores can be attributed to the treatment of two types of feedback.

To investigate the learning gains for each error category, the individual scores on sentence fragments, run-ons, and subordinate clause errors in the pre- and post-tests are displayed in Table 4.8, together with the pre-to-post-test gain scores in each category. In addition, Figure 4.11 provides a visual representation of the mean scores of each error type in the pre- and post-tests.

Table 4.8 Descriptive statistics of participants’ scores for correcting all three types of syntactic errors and participants’ gain scores on pre- and post-tests ($N_{\text{corpus}} = 45; N_{\text{code}} = 45$)

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>Error type</th>
<th>FRAG Mean (out of 7)</th>
<th>SD</th>
<th>RO Mean (out of 7)</th>
<th>SD</th>
<th>CL Mean (out of 7)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-</td>
<td>Corpus</td>
<td></td>
<td>1.14</td>
<td>1.42</td>
<td>1.19</td>
<td>1.74</td>
<td>1.59</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td></td>
<td>1.21</td>
<td>1.55</td>
<td>1.12</td>
<td>1.35</td>
<td>1.53</td>
<td>1.31</td>
</tr>
<tr>
<td>Post-</td>
<td>Corpus</td>
<td></td>
<td>3.68</td>
<td>1.78</td>
<td>4.61</td>
<td>1.95</td>
<td>3.05</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td></td>
<td>2.20</td>
<td>1.94</td>
<td>2.43</td>
<td>2.02</td>
<td>2.10</td>
<td>1.76</td>
</tr>
</tbody>
</table>
Notes: The total score for each error category is 7.

The data in Table 4.8 and the plot lines in Figure 4.11 show that the two groups had similar mean pre-test scores in each error category (Frag: $M_{\text{corpus}}$ vs $M_{\text{code}}$: 1.14 vs 1.21; Ro: $M_{\text{corpus}}$ vs $M_{\text{code}}$: 1.19 vs 1.12; Cl: $M_{\text{corpus}}$ vs $M_{\text{code}}$: 1.59 vs 1.53). However, students from the corpus group had higher mean post-test scores when correcting sentence fragments, run-on sentences, and subordinate clause errors ($M = 3.68$, $SD = 1.78$; $M = 4.61$, $SD = 1.95$; and $M = 3.05$, $SD = 1.65$, respectively) than those in the code group ($M = 2.20$, $SD = 1.94$; $M = 2.43$, $SD = 2.02$; and $M = 2.10$, $SD = 1.76$). Students in both groups had the most improvement when correcting run-on sentences ($M_{\text{corpus}} = 3.42$; $M_{\text{code}} = 1.31$) and the least improvement when correcting subordinate clause errors ($M_{\text{corpus}} = 1.56$; $M_{\text{code}} = .57$).

Figure 4.10 Mean scores for correcting all three types of syntactic errors on pre- and post-tests ($N_{\text{corpus}} = 45$; $N_{\text{code}} = 45$).
4.3.2. Qualitative Data Analysis

In addition to the quantitative results, the qualitative data from the semi-structured interviews clarified the reasons for students’ successful or unsuccessful corrections in the post-test. During the interview, students revisited their pre-test and post-test papers. They were asked to delineate reasons for their successful and unsuccessful experiences correcting errors under two circumstances: 1) successful correction of the target error in the post-test and unsuccessful correction of an error in the pre-test, and 2) unsuccessful correction of the target error in both the pre-test and post-test.

4.3.2.1. Successful correction of errors in the post-test and unsuccessful correction/n of errors in the pre-test

In the case of both groups, students mainly attributed their successful correction of errors in the post-test compared to unsuccessful error correction in the pre-test to their interaction with the feedback during the previous exercises (Table 4.9).

<table>
<thead>
<tr>
<th>Reasons for successful error correction</th>
<th>Corpus group</th>
<th>Code group</th>
</tr>
</thead>
<tbody>
<tr>
<td>N= 9 (27 units)</td>
<td></td>
<td>N= 9 (30 units)</td>
</tr>
<tr>
<td>1</td>
<td>Learning from the corpus-based feedback 77.78% (21 units)</td>
<td>Learning from the coded feedback 73.33% (22 units)</td>
</tr>
<tr>
<td>2</td>
<td>22.22% (6 units)</td>
<td>Having known already how to correct target errors, and just skipping errors in the pre-test 16.67% (5 units)</td>
</tr>
<tr>
<td>3</td>
<td>Practicing in personal writing 10% (3 unit)</td>
<td></td>
</tr>
</tbody>
</table>
Learning from the feedback. People in both groups felt that they had learned syntactic knowledge from the feedback on the error correction exercises. A few examples of students’ utterances are cited below:

a. I think that I learned that from the feedback. They marked errors in previous assignments [error correction practices], and when I see them again, I can recognize and correct. (Student 8, high-improvement learner, comment on a subordinate clause error, corpus group)

b. Something from in the feedback. It’s always helpful. (Student 5, intermediate-improvement learner; comment on a sentence fragment, corpus group)

c. I’ve learned how to correct it. I always need to add “and,” “but”, “so” between two sentences. So I changed the sentence and make them connected with...“and.” (Student 4, intermediate-improvement learner; comment on a sentence fragment, corpus group)

d. Yeah, I think in the pretest I really didn’t recognize any clauses. So yeah, the exercises and the feedback, like examples, were really helpful to me. (Student 18, high-improvement learner, comment on a subordinate clause error, code group)

e. I remember that from the feedback in previous tasks [error correction exercises]. (Student 16, high-improvement learner; comment on a sentence fragment, code group)

Having already known how to correct target errors. In addition, 22.22 % of students from the corpus group and 16.67 % from the code group said that they knew how to correct target errors already in the pre-test, and they just skipped the errors.

a. Oh, this one. I should have corrected it. I guess I just skipped it. (Student 9, high-improvement learner; comment on a subordinate clause error, corpus group)

b. I think I know how to correct it in the pre-[test]. I am not sure why I didn’t. (Student 11, low-improvement learner; comment on a sentence fragment, code group)

Practicing in personal writing. One student in the code group remarked that he utilized knowledge learned from the coded feedback to correct his own errors in his own
writing. Such practice raised his error awareness and empowered him to correct errors in new texts.

a. I guess the practice makes me improve. When I learn something, I want to use it in my life and then I learn gradually. (Student 18, high-improvement learner; comment on a subordinate clause, code group)

b. I always write run-on sentences, so after the practice, I pay attention to those errors in my writing. So in the post-test, I can correct them. That helped me learn. (Student 18, high-improvement learner; comment on a run-on sentence, code group)

4.3.2.2. Unsuccessful correction of target errors in the post-and the pre-test

When students were unable to correct target errors in the pre-tests and unable to correct target errors again in the post-test after the error correction practice, they were asked about the reasons behind their unsuccessful learning experience. The four main reasons are listed in Table 4.10, and each reason is elaborated below with students’ comments from the semi-structured interviews.

Table 4.10 Reasons for unsuccessful error correction in the post-test.

<table>
<thead>
<tr>
<th>Reasons for unsuccessful error correction</th>
<th>Corpus group</th>
<th>Code group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure of error noticing</td>
<td>51.85% (14 units)</td>
<td>Not learning with the feedback or incomprehension of the feedback 41.38% (12 units)</td>
</tr>
<tr>
<td>Forgetting what had been learned with the feedback</td>
<td>29.62% (8 units)</td>
<td>Failure of error noticing 37.93% (11 units)</td>
</tr>
<tr>
<td>The length and complexity of sentence structures in the new article</td>
<td>11.11% (3 units)</td>
<td>The length and complexity of sentence structures in the new article 10.34% (3 units)</td>
</tr>
</tbody>
</table>
Failure of error noticing. Skipping errors was the most common reason given by students from the corpus group (51.85% of the idea units) and the second-most-common reason from the code group (41.38% of the idea units) to explain their failure to correct target errors on the post-test. Interviewees from both groups explained that they intuitively felt the sentences with errors were correct; however, they felt that if the target sentence was highlighted as incorrect, or if they read the sentence again, they would find and correct the errors.

a. I know how to correct it. This is what I learned from the feedback, but I didn’t... [Researcher: Then could you think about why you didn’t correct it in the post-test?] Probably the sentence isn’t highlighted. If the sentence is highlighted as incorrect, I can correct. (Student 4, intermediate-improvement learner; comment on a subordinate clause, corpus group)

b. The error is not very obvious. [Researcher: Do you think you know how to correct it?] When I read it again, I will delete “there are.” [Researcher: So why didn’t you find it in the test?] It’s hard. When I read, I read silently... it [the sentence] just sounds too right. But if I read it now, I will notice the mistake. (Student 8, high-improvement learner; comment on a subordinate clause, corpus group)

c. If this sentence is marked as incorrect, I can find it and correct... like deleting “in” here, but if it’s not marked, I think it’s right. [Researcher: Why?] It sounds correct, and I just leave it there. (Student 5, intermediate learner, corpus group; comment on sentence fragments, corpus group)

d. Yeah, this one is incorrect. I need to add changed “comma” to “period” here and change “comma” to “period” here. [Researcher: Then why do you think you didn’t correct it in the post-test?] I guess I didn’t find it in the [post-]test.... If I read it again and again, I may find it, but it sounds correct. (Student 14, intermediate-improvement learner; comment on a run-on sentence, code group)
e. I think I just missed it [in the post-test]. [Researcher: Then why didn’t you catch it?] Maybe I need some hints. (Student 13, intermediate-improvement learner; comment on a run-on sentence, code group)

**Forgetting what had been learned with the feedback.** Additionally, students from both groups (29.62% from the corpus group and 10.34% from the code group) pointed to forgetting what had been learned from the feedback as another factor that contributed to their unsuccessful knowledge learning. Corpus group students attributed their forgetfulness to the limited exposure to the feedback and limited numbers of practice.

a. It has been for a while, for a couple of days. I did not listen or rerevise based on the example. So I forgot [how to correct it]. [Researcher: Do you think the feedback is helpful?] I think the feedback is helpful, but I think students should keep revise and practice... Students [who] can practice more can have a better result. (Student 9, intermediate-improvement learner, corpus group; comment on a subordinate clause error)

b. I forgot it totally... I mean I corrected it in the exercises, but I probably need more practice. (Student 8, high-improvement learner; comment on a sentence fragment, corpus group)

c. No one can remember everything. So repetition is important. If we practice this more using the feedback, we may remember. It's only a process. (Student 1, low-improvement learner; comment on a subordinate clause error, corpus group)

d. Oh, I made it correct previously [in error correction exercises]. I can’t remember it. (Student 16, high-improvement learner; comment on a sentence fragment, code group)

e. I guess I forgot the feedback. [The students looked at their corrections in the error correction exercises.] I corrected it in here, but I forgot it now. (Student 18, high-improvement learner; comment on a subordinate clause error, code group)

**The complexity of the sentence structures in the new article.** The complexity of the sentence structures and unknown words in the new article was third factor mentioned by interviewees from both groups (11.11% from the corpus group and 10.34% from the code group). Students failed to correct the sentence, either because they were afraid of
changing the original meanings of the long sentences, or because they were unable to deal with errors in complex sentences.

a. Basically, just too many commas and these are almost interconnected and I cannot separate it, and if I did I would lose the meaning. It’s one of those sentences where you know it’s a long sentence, and [you could] break it, but it’s very hard. Very difficult to break it. (Student 8, high-improvement learner; comment on a run-on sentence, corpus group)

b. It’s very long and has many clauses and clauses. I want to add something here, like “is,” but if I changed it, the meaning will change too. So I leave it as original. (Student 16, high-improvement learner; comment on a run-on sentence, code group)

c. The sentence is kind of complicated, so I don’t know how to deal with it. (Student 12, low-improvement learner; comment on a subordinate clause, code group)

Not learning with the feedback or incomprehension of the feedback. Students in both groups found that their failure to learn with the feedback or to understand the feedback also accounted for their unsuccessful corrections in the post-test. For example, one student in the corpus group said that she failed to correct a run-on sentence because of her incomplete understanding of the feedback. She felt that even though she corrected errors using feedback in the error correction exercises, she was unable to comprehend the reasons for the errors and ways of correcting them given in the feedback. Such incomprehension resulted in her inability to transfer knowledge to correct errors in other contexts:

I only know how to correct a run-on error with the comment, but don’t know why, I think. If I don’t understand it [the feedback] completely, I can’t use it here [in the post-test]. That’s the reason I think. (Student 3, low-improvement learner; comment on a run-on sentence, corpus group)

In the code group, the failure of learning with the feedback ranked as the primary reason for the unsuccessful correction of errors in the post-test. Seven students in this group felt that insufficient or unclear information from the coded feedback hindered their
understanding of the errors, thus leading to ineffective learning and a lack of awareness of errors in the subsequent texts. The following comments illustrate students’ opinions.

a. I feel the feedback doesn’t help me learn clause [errors]. Like here, I can’t find errors with the feedback [in error correction exercises]. [Researcher: Why?] I was not learning because “cl” is not clear. I don’t know why it is a “cl” error where it is, and how to correct it. So I can’t correct it in the paper [in the post-test]. (Student 11, low-improvement learner; comment on a subordinate clause error, code group)

b. Actually, I didn’t get it in the exercises. So I can’t correct it here [in the post-test]. (Student 10, low-improvement learner; comment on a sentence fragment, code group)

c. I didn’t notice it [the error in the post-test]... even now you pointed it to me. [Researcher: Did you learn how to correct it from the feedback and exercises?] ...Not really. (Student 18, high-improvement learner; comment on a subordinate clause error, code group)

Two students also pointed out that even though coded feedback helped raise their awareness of target errors in new texts, they did not learn how to correct errors and were unable to correct the errors by themselves.

a. I know the sentence sounds weird, but I am not sure where exactly the error is... [Researcher: Do you think you learned anything from the feedback?] I think so, so I can find it, but I’m not sure how to correct it. Student 13, intermediate-improvement learner; comment on a subordinate clause error, code group)

b. I know this is wrong [in the post-test] because I learned it from the exercises [with the feedback]. I didn’t correct because the information doesn’t teach me the errors and ways of correction. (Student 11, low-improvement learner; comment on a sentence fragment, code group)

The four main reasons for unsuccessful error correction in the post-test unveiled key factors that affect the syntactic knowledge learning and knowledge transfer from error correction exercises to the post-test. First, according to the Interactionist Approach, comprehensible input is crucial for second language learning. When discussing the effectiveness of the feedback on noticing and correcting errors in error correction
exercises, nine participants from the corpus group generally found that corpus-based feedback was comprehensible and helpful, while most participants from the code group thought the coded feedback lacked sufficient information to help understand errors. Unsurprisingly, the lack of comprehension of the feedback was mentioned as the most common reason for unsuccessful error correction by the code group, but was the least common reason given by the corpus group when reflecting on the unsuccessful error correction in the post-test. The qualitative data from think-aloud protocols and semi-structured interviews indicated that the comprehensibility of the feedback is a key factor that contributes to language learning. In addition to the quality of the feedback, the repeated practices and repeated exposure to the feedback are other factors that assist in the language learning. Participants from both groups said that they failed to notice errors in a new text and forgot what they had learned from the feedback two weeks after the error correction exercises. They found that more practice and more interaction with the feedback are needed to reinforce and automate their language knowledge. Such statements echo the Skill-learning theory that students develop procedural knowledge via repeated practice. Thirdly, the complexities of sentences in the new text may have also affected students’ knowledge application and transfer. Students said that they hesitated to make corrections when the sentences in the new text were long and complicated.

After exploring the reasons for their successful and unsuccessful knowledge learning of syntactic errors, students from the two groups voted on the most and least difficult error types in the post-test. Subordinate clause errors were perceived as the most difficult to correct and run-on sentences the least difficult in a new article without feedback, as seen in Chart 4.12. Such claims were substantiated by the pre- and post-test
results, in which gain scores from pre- to post-tests of subordinate clause errors were the lowest and run-on sentences were the highest in the two groups. The following utterances from both groups illustrate the reasons.

![Most difficult Syntactic Errors for Correction in the Post-test](image1)

**Figure 4.11** *The most and least difficult errors for correction in the post-test (N corpus interviewee = 9; N code interviewee = 9).*

**Clause errors.** Both the corpus and code groups found subordinate clause errors to be the most challenging to correct in the post-test:

a. [Researcher: Which error types do you think are more difficult to be corrected without the feedback?] Clause. The structure of the sentence can be different with different errors. (Student 2, low-improvement learner, comment on subordinate clause errors, corpus group)

b. I think clause errors are more difficult [than sentence fragments and run-on sentences]. Like fragments, I know they may miss a subject or verb or something else. Run-on sentences need to be separated. But clause errors are complicated. It may need to add something, delete something, or change the structures. So it’s very difficult to correct it. (Student 14, intermediate-improvement learner; comment on subordinate clause errors, code group)

**Run-on sentences.** Both groups unanimously voted that run-on sentences were the easiest to correct in a new text after receiving feedback.

a. Run-ons is easy than fragment and... clause. I didn’t know [in the pre-test], I didn’t know it’s a run-on. After learning this, I can use it in different situations. The rules are straightforward... I would say “it should be a period.” (Student 5, intermediate-improvement learner; comment on run-on sentences, corpus group)
b. It's [run-on sentences] easier than others [sentence fragments and subordinate clause errors]. It is more understandable why it’s wrong. (Student 17, high-improvement learner; comment on run-on sentences, code group)

The qualitative data in this section illustrate the reasons for students’ successful and unsuccessful application of knowledge to a new text. Complementing the quantitative findings, the qualitative data explain why students in both groups made the greatest gains in correcting run-on sentences and the lowest gains correcting subordinate clause errors.

4.3.3. Triangulation of Quantitative and Qualitative Results

Triangulation of the data reveals the language gains for each group following the practice with the target feedback and the reasons for students’ successful and unsuccessful knowledge learning.

The quantitative analysis of students' improvement from the pre- to post-test produced evidence showing students' learning gains. Conducting a two-way mixed ANOVA, the analyses indicated that, although students from both groups had significant improvement on their syntactic knowledge from the pre-test to the post-test, the learning gains in the corpus group were statistically significantly higher than those in the code group. What is more, the descriptive statistics show that both groups had similar mean pre-test scores for sentence fragments, run-on sentences, and subordinate clause errors, while the corpus group’s post-test mean scores for each error category were higher than those from the code group.

The qualitative data from the semi-structured interviews provided insight into the reasons for students’ successful and unsuccessful learning experiences. When discussing their successful error corrections on the post-test, students in both groups felt that they
had learned from the target feedback and were able to apply that knowledge to the new text. When discussing their unsuccessful error correction experience on the post-test, students in both groups articulated four main reasons behind their lack of success: failure to notice errors, forgetting what had been learned with the feedback, the length and complexity of sentence structures in the new article, and not learning with the feedback. The primary reason for unsuccessful correction given by the corpus group was failure in error noticing: students already possessed the knowledge for error correction, but they needed hints or more practice to increase their ability to notice errors. By contrast, in the code group, a lack of knowledge about correcting syntactic errors was the main reason given to explain students’ unsuccessful error correction on the post-test: students said that they did not receive sufficient information for error correction from the feedback. In terms of correction for each error category on the post-test, both groups unanimously agreed that subordinate clause errors constituted the most difficult error for correction, and run-on sentences the easiest. They felt that subordinate clause errors, found in complex sentences, were difficult to notice and correct; students in both groups desired more examples and explanations to illustrate clause errors. On this point, students’ observations were consistent with their performance on the post-test, as both groups showed the least amount of improvement on subordinate clauses and the most on correcting run-on sentences.

4.4. Chapter Summary

This chapter discussed the findings in response to the three research questions concerning the language learning potential of two types of CALL corrective feedback: corpus-based feedback and coded feedback. The quantitative findings suggested that
corpus-based feedback helped students notice and correct syntactic errors in error
correction practice exercises more than coded feedback did. Moreover, corpus-based
feedback was more beneficial for facilitating syntactic knowledge application to a new text
in the post-test, despite the fact that both groups improved significantly from the pre- to
post-test. The corpus group’s error noticing scores on the error correction practice
exercises and their pre- to post-test gain scores for each error category—sentence
fragments, run-on sentences, and subordinate clause errors—were also higher than those
of the code group.

The think-aloud protocols and semi-structured interviews revealed students’
perceptions of the two types of feedback. Discussing the first two aspects of language
learning potential, 96.48 % of remarks from interviewees from the corpus group about
corpus-based feedback for error noticing were positive, as were 87.78 % of their comments
on error correction. Students from this group emphasized that various features of the
corpus-based feedback, such as colors, examples, and oral explanations, facilitated their
noticing of errors and enhanced their understanding. Nonetheless, 3.52 % of the comments
on error noticing were negative, as were 12.22 % of the utterances concerning error
correction. These students found the oral explanations in the corpus-based feedback
unnecessary for encouraging error noticing, and they remarked on the need for more
examples and explanations to facilitate their understanding. By comparison, 58.43 % of the
comments on the coded feedback were positive in regard to error noticing, as were
46.88 % of the comments on error correction. Interviewees said the error codes informed
them of error types, facilitating their ability to search for and correct errors. However,
41.57 % of the comments on the coded feedback for error noticing were negative, as were
53.12% of the comments on error correction. Students in the code group cited the lack of examples and explanations as the primary reason for their failure to notice errors and make appropriate corrections. Regarding the third aspect of language learning, knowledge learning application, both groups reported experiencing knowledge development with the target feedback. When discussing unsuccessful knowledge application experience, students in the corpus group said they needed more hints to find and correct errors, while the code group said that the feedback failed to equip them with enough knowledge to enable knowledge learning and error correction on the post-test.

Based on the Logical Framework and its evaluation criteria, the next chapter continues the evaluation of the two types of corrective feedback by examining two other qualities: learner fit and impact.
CHAPTER 5. RESULTS ON THE LEARNER FIT AND IMPACT OF CALL CORRECTIVE FEEDBACK

This chapter discusses the results pertaining to two qualities of CALL corrective feedback: learner fit and impact. According to Chapelle (2001), “learner fit” refers to the appropriateness of CALL materials for facilitating learning in the target group of students. “Impact” refers to the effects students experience from interacting with CALL materials. The examination of each of these qualities was performed by analyzing students’ responses to items on a seven-point Likert-Scale questionnaire and to questions posed in semi-structured interviews.

5.1. Learner Fit of Corpus-based and Coded Feedback

Based on the logical framework for the development and evaluation of CALL corrective feedback, the criterion used to investigate the hypothesis concerning learner perceptions of the appropriateness of the feedback is learner fit. This analysis addresses the fourth research question:

- What evidence suggests that advanced-low ESL learners perceive that corpus-based feedback is more appropriate than coded feedback for enabling correction of syntactic errors?

This question was addressed through data collected from students’ responses to nine questionnaire items and related semi-structured interview questions. The quantitative and qualitative data were juxtaposed to illustrate whether and how the two types of corrective feedback fit students’ needs and facilitated individualized learning. Overall, the data yielded strong evidence that advanced-low ESL learners found corpus-
based feedback to be more appropriate than coded feedback for enabling correction of syntactic errors.

5.1.1. Questionnaire Data

Questionnaire data on Learner Fit for each type of corrective feedback were collected from 90 participants’ responses to nine seven-point Likert-Scale items measuring three sub-constructs: 1) clarity of feedback, 2) usefulness of feedback, and 3) self-correction with feedback. Two negatively worded items (see Table 5.1) were used in the questionnaire to determine whether each participant fully understood those items and responded to each earnestly. Four students from the corpus group and three from the code group consistently responded with the same option regardless of the item’s statement; consequently, their answers were cleaned from the data set. In total, the responses from 83 participants (41 from the corpus group and 42 from the code group) were used for data analysis. The negatively worded survey items were reverse-coded in order to create a meaningful total score when combined with the positively worded items. Table 5.1 catalogs the questionnaire items and presents the analytical results, including the mean and standard deviation in each group.
Table 5.1 Questionnaire items on learner fit and participants’ responses to the Likert-Scale items in the questionnaire

<table>
<thead>
<tr>
<th>No.</th>
<th>Questionnaire items</th>
<th>Corpus group N=41</th>
<th>Code group N=42</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>Clarity of feedback</td>
<td>6.03</td>
<td>.84</td>
</tr>
<tr>
<td>1.</td>
<td>The feedback was understandable in the error correction practices.</td>
<td>5.88</td>
<td>1.08</td>
</tr>
<tr>
<td>2.</td>
<td>The feedback helped to explain the errors clearly.</td>
<td>6.05</td>
<td>.92</td>
</tr>
<tr>
<td>3.</td>
<td>I had difficulty understanding the feedback. (reverse coding for analysis)</td>
<td>6.17</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>Usefulness of feedback</td>
<td>6.06</td>
<td>.85</td>
</tr>
<tr>
<td>4.</td>
<td>The feedback was useful for teaching me how to correct the mistakes in the error</td>
<td>5.98</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>correction exercises.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>The feedback was helpful for correcting sentence fragments, run-ons, and clause</td>
<td>6.07</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>mistakes in the error correction exercises.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>I found the feedback useless in helping me correct sentence fragments, run-ons, and</td>
<td>6.15</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>clause mistakes in the error correction exercises. (reverse coding for analysis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-correction with feedback</td>
<td>5.69</td>
<td>.83</td>
</tr>
<tr>
<td>7.</td>
<td>The feedback helped me revise the error correction practices by myself.</td>
<td>5.66</td>
<td>1.13</td>
</tr>
<tr>
<td>8.</td>
<td>I could use the feedback to correct sentence fragments, run-ons, and clause</td>
<td>5.71</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>mistakes in the error correction exercises without any additional help.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>The feedback itself was enough to help me revise mistakes in the error correction</td>
<td>5.71</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>exercises on my own.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Cronbach's alpha of the nine Likert-scale questions was .862, indicating that the questions largely measure the same construct. The nine questionnaire items were
employed to measure three sub-constructs of learner fit quality, with three items associated with each sub-construct. The first sub-construct was “clarity of feedback.” Item three under this construct, which was negatively worded, was reversed for data analysis. The scale had a high level of internal consistency, as determined by a Cronbach’s alpha of .838. The second sub-construct, “usefulness of feedback,” also comprised three items. Again, the third was negatively worded, and students’ responses to this question were reversed for data analysis. The Cronbach’s alpha of these three items was .910. The third sub-construct was “self-correction with the feedback.” The scale had an acceptable level of internal consistency, as determined by a Cronbach’s alpha of 0.781.

The descriptive data for the total score of each sub-construct are summarized in Table 5.2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Questionnaire item</th>
<th>Corpus group</th>
<th>Code group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N = 41</td>
<td>N = 42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total  SD</td>
<td>Total  SD</td>
</tr>
<tr>
<td>1.</td>
<td>Clarity of the feedback</td>
<td>18.10 2.53</td>
<td>15.07 3.06</td>
</tr>
<tr>
<td>2.</td>
<td>Usefulness of the feedback</td>
<td>18.20 2.56</td>
<td>14.83 3.16</td>
</tr>
<tr>
<td>3.</td>
<td>Self-correction with feedback</td>
<td>17.07 2.49</td>
<td>14.69 3.15</td>
</tr>
</tbody>
</table>

As seen in the table, the total scores of the items measuring clarity of feedback, usefulness of feedback, and self-correction with feedback were higher in the corpus group (Total corpus 1 = 18.10, SD corpus 1 = 2.53; Total corpus 2 = 2.56 18.20, SD corpus 2 = 2.56; Total corpus 3 = 17.07, SD corpus 3 = 2.50) than the code group (Total corpus 1 = 15.07, SD corpus 1 = 3.06; Total corpus 2 = 14.83, SD corpus 2 = 3.16; Total corpus 3 = 14.69, SD corpus 3 = 3.15).

A one-way ANOVA was conducted to determine whether students’ perceptions of learner fit for each type of corrective feedback were significantly different from each other.
As assessed by inspection of the boxplot in Figure 5.1, there were outliers among the total scores of each sub-construct of the corpus group, as assessed by the inspection of the boxplot in Figure 5.1.

![Boxplot of total scores of each construct concerning the learner fit of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21).](image)

Figure 5.1 *Boxplot of total scores of each construct concerning the learner fit of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21).*

The results from the original data with and without outliers were reported and compared. With the original data, the total scores of the corpus group’s assessment regarding clarity of feedback ($p = .001 < .05$), usefulness of feedback ($p = .001 < .05$), and self-correction with feedback ($p = .024 < .05$) were not normally distributed when assessed by Shapiro-Wilk’s test. The total scores of the code group for clarity of feedback ($p = .113 > .05$), and usefulness of feedback ($p = .662 > .05$) were normally distributed, but those for self-correction with feedback ($p = .024 < .05$) were not. The assumption of homogeneity of variances for clarity of feedback ($P = .063 > .005$), usefulness of feedback ($P = .068 > .005$), and
and self-correction with feedback (P = 0.094 > 0.005), as assessed by Levene’s test of the homogeneity of variances. The results with the one-way ANOVA showed that the scores for each sub-construct—clarity of feedback (F (1, 81) = 24.11, p < 0.0005, partial η2 = 0.23), usefulness of feedback (F (1, 81) = 28.25, p < 0.0005, partial η2 = 0.26), and self-correction with feedback (F (1, 81) = 14.55, p < 0.0005, partial η2 = 0.15)—were significantly different. The effect sizes indicated 23%, 26% and 15% of variance in questionnaire scores of the clarity of feedback, usefulness of feedback, and self-correction with the feedback can be attributed to feedback types.

When the outliers were deleted from the original data, there were no new outliers in the data set. The data of the corpus group in the new set regarding self-correction (p = 0.060 < 0.05) were normally distributed, while clarity of feedback (p = 0.016 < 0.05), usefulness of feedback (p = 0.006 < 0.05) were not, when assessed with Shapiro-Wilk’s test. The total scores of the code group on clarity of feedback (p = 0.113 > 0.05), usefulness of feedback (p = 0.662 > 0.05), and self-correction with feedback (p = 0.339 > 0.05) were not normally distributed. The assumption of homogeneity of variances for each construct were not met, respectively (pclarity = 0.13; pusefulness = 0.04; pself-correction = 0.24), as assessed by Levene’s test of the homogeneity of variances. The results showed that the scores for each sub-construct—clarity of feedback (Welch’s F (1, 73.27) = 28.09, p < 0.0005, partial η2 = 0.27), usefulness of the feedback (Welch’s F (1, 69.89) = 40.90, p < 0.0005, partial η2 = 0.32), and self-correction with feedback (Welch’s F (1, 73.16) = 19.13, p < 0.0005, partial η2 = 0.18)—were significantly different. The effect sizes indicated 27%, 32% and 18% of variance in questionnaire scores of the clarity of feedback, usefulness of feedback, and self-correction with the feedback can be attributed to feedback types. In summary, the results of the one-
way ANOVAs indicated that the corpus group had more positive perceptions about the appropriateness of the corpus-based feedback for enabling correction of syntactic errors compared to the code group. As shown in the following section, students’ utterances in the semi-structured interviews further illustrated their perceptions of the learner fit quality of the two types of corrective feedback.

5.2.2. Semi-structured Interview Data Concerning Impact

The analysis of interview data revealed 55 idea units related to three aspects of learner fit, as shown in Table 5.3: clarity of feedback, appropriateness of feedback for task completion, and suitability of feedback for individual learner characteristics. The following sections discuss each sub-theme concerning learner fit and provide representative quotations from the interviews.

Table 5.3 Positive and mixed/negative idea units identified with each theme and subtheme regarding learner fit, as reported in semi-structured interview transcripts

<table>
<thead>
<tr>
<th>Theme</th>
<th>Sub-theme</th>
<th>Corpus group N =9</th>
<th>Code group N =9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Mixed/Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>Clarity of feedback</td>
<td>89.29% (25 units)</td>
<td>10.71% (3 units)</td>
<td>29.63% (8 units)</td>
</tr>
<tr>
<td>Appropriateness of feedback for task completion</td>
<td>32.14% (9 units)</td>
<td>0% (0 units)</td>
<td>11.11% (3 units)</td>
</tr>
<tr>
<td>Suitability for individual learner characteristics</td>
<td>25.00% (7 units)</td>
<td>3.7% (1 unit)</td>
<td>11.11% (3 units)</td>
</tr>
<tr>
<td>Suitability to individual learning pace</td>
<td>7.14% (2 units)</td>
<td>0% (0 units)</td>
<td>0% (0 units)</td>
</tr>
</tbody>
</table>
Clarity of feedback. Perceptions of the clarity of the feedback varied between the two groups. 89.29% of the utterances from the corpus group described the feedback as easy to understand, while only 29.63% of the utterances from the code group reflected this position. Specifically, seven students in the corpus group noted that they had no difficulty understanding the feedback. Participants’ ease in understanding the feedback is expressed in the following comments:

a. Yes, I can understand it pretty well. The words are easy and the explanations are good. (Student 3, low-improvement learner, corpus group)

b. Yes, I can understand them. (Student 6, intermediate-improvement learner, corpus group)

c. I think the feedback is understandable, especially when I compare them, the incorrect and correct examples, to mine [incorrect sentences in error correction practice exercises], it’s nice and clear. (Student 9, high-improvement learner, corpus group)

Two interviewees in the corpus group expressed ambivalence. One student (Student 8, high-improvement learner, corpus group) found certain grammatical terms difficult to understand and would have preferred more explanation. She commented, “Yes, I can understand most of them, but as I mentioned, some grammar terms are confusing sometimes. If there are some explanations include in the feedback, I will learn better.”

Furthermore, another student (Student 7, high-improvement learner, corpus group) would have preferred the feedback to include more example sentences to facilitate her understanding. She said, “I can understand those feedback, but I guess more examples are helpful. I wanted to see more and understand it better.”

Interviewees from the code group also commented on the clarity of the feedback in the interview. Two students found the feedback fully understandable:
a. The codes are good [for understanding]. (Student 17, high-improvement learner, code group)

b. I think I can [understand]. They are simple, and when I looked at the codes and their definitions in the sheet, I can understand. (Student 18, high-improvement learner, code group)

At the same time, the remaining seven interviewees made negative or mixed comments about their understanding of the coded feedback. Three interviewees found the coded feedback too abstract to understand:

a. I don’t understand the feedback. They are very vague to me. I think more explanations were needed. (Student 11, low-improvement learner, code group)

b. I always tried to correct, but I don’t really understand the feedback too much. I think I just correct it by myself. (Student 13, intermediate-improvement learner, code group)

c. The feedback is too abstract. I had a hard time [understanding them]. (Student 10, low-improvement learner, code group)

Four found they could understand some coded feedback on specific types of syntactic errors, but not all of it:

a. Some of them are good, but not all of them. For example, I know sentence fragment, so I can understand the feedback. I had no ideas about clause errors. When the feedback said “Cl,” I got lost. (Student 14, intermediate-improvement learner, code group)

b. I can’t say I can’t understand all of them, but probably 60% hard to understand. Especially, some fragments and clauses, I am wondering why they are errors, cause I can’t find errors, or I found it’s weird, but doesn’t sound like a fragment or clause [error]. (Student 15, intermediate-improvement learner, code group)

c. I understand fragments and clause sentences, but run-on sentences are hard. (Student 16, intermediate-improvement learner, code group)

**Appropriateness of feedback for task completion.** Interviewees also reflected on their perception of the appropriateness of the feedback for completing the error correction
practice exercises. All students in the corpus group found the feedback effective in helping them locate and correct errors in the exercises.

a. Yes, it teaches me a lot. I can notice errors with colors, examples in the feedback so I can finish the task in the class. (Student 1, low-improvement learner, corpus group)

b. Of course. Some errors are difficult to find. Like sentence fragments missing something or clause errors having too much clauses together. They are very complex, but the feedback really help find and correct them. (Student 5, intermediate-improvement learner, corpus group)

c. Yes, I will say it definitely help me a lot with the task. It provided a lot of useful information to me, helped me better understand the errors, and explained very clearly how to correct the errors. (Student 8, high-improvement learner, corpus group)

Interviewees in the code group, however, reported a different experience. Three students had a positive perception of the coded feedback for task completion. Among those three, two said that they already possessed syntactic knowledge about the errors; in these cases, the feedback was useful in reminding them of their pre-existing knowledge and facilitating completion of the task. The third student found that the feedback allowed him to contemplate the errors and learn in the process.

a. The feedback is ok for the task. Cause I know those errors, but sometimes couldn’t find them, so the feedback helped me. (Student 17, high-improvement learner, code group)

b. Yes, it’s appropriate. I feel that those errors have been taught in the high school or when I prepared exams, so the feedback reminds me. (Student 18, high-improvement learner, code group)

c. I think it is [appropriate]. I learned in this process and the feedback gives me clues, and I don’t 100% depend on the feedback. When I keep thinking, I can find the inappropriate part of the sentences, and tried to fix that. (Student 14, intermediate-improvement learner, code group)
However, six other students had either negative or mixed opinions about the appropriateness of the feedback for error correction tasks. Three felt it was very difficult to complete the task with the coded feedback. The other three said that the feedback helped them with the task to some extent, but they felt the help they received was insufficient for correcting some errors.

a. No. It doesn’t include any explanations I need, it doesn’t help me complete the task, so it’s not appropriate for me. (Student 12, low-improvement learner, code group)

b. I feel that feedback itself is difficult for me to use [in error correction practice exercises]. For example, the feedback tells me there is a clause error in the sentence. Many times I cannot find it, and cannot correct it. (Student 13, intermediate-improvement learner, code group)

c. I think it may not. If the task asked us to correct errors, like spelling errors or verb errors, I think the feedback is good for me. For the task asked to correct errors like fragments, I think they are harder and the feedback should provide more than this. (Student 15, intermediate-improvement learner, code group)

**Suitability for individual learner characteristics.** Students in both groups evaluated the feedback in terms of its suitability to their individual needs. Six interviewees from the corpus group discussed a variety of features that suited their own learning styles:

a. Yeah. I am an engineer student, so I want to know reasons [of the errors] in a direct way. So I found the rule section is clear. I take a look at that part first. If I can’t get it, I will read examples and explanations. (Student 2, low-improvement learner, corpus group)

b. I feel that I like the feedback meets my needs. I like listening, but I don’t usually get spoken feedback, so this time, I have choices. (Student 4, intermediate-improvement learner, corpus group)

c. I couldn’t ask more. I like examples, colors, oral part. So I used the examples to learn how to correct errors and then I listen to oral feedback to make sure I understand it and learn further with it. (Student 6, intermediate-improvement learner, corpus group)
Two students also described how the feedback helped them maintain their learning pace. They felt that the feedback allowed them the freedom to maintain their personal learning style.

a. I don't learn quickly. I need to read and think, and sometimes to read again. The feedback helped me do my style. (Student 1, low-improvement learner, corpus group)

b. And I liked the feedback cause I can choose whether I read the feedback or listen. If I was required to listen to every feedback, I will be annoyed, but this feedback allows me to choose. If I need, I listen. If not, I will move to next one. (Student 5, intermediate-improvement learner, corpus group)

One student (Student 7, high-improvement learner, corpus group) expressed her need for more examples in the feedback. She said she preferred reading and learning with examples; therefore, including more examples in the feedback would better serve her learning style.

Students in the code group felt differently about the individual appropriateness of the feedback. Three students expressed their sense that the feedback met their individual needs and learning styles.

a. I still liked the feedback. I don't think the feedback can provide me everything. Like in each assignment [in a writing course], I got such feedback. If I don't understand, I will search key words on the internet. Feedback already told me sentences are wrong and error name, so I can work out by myself. (Student 12, low-improvement learner, code group)

b. Yes, I can correct those errors by myself, but sometimes, I just missed them. The feedback really helped me to notice errors and gives me clues. (Student 14, intermediate-improvement learner, code group)

c. Yes, for me, it teaches me what the error is and I can figure out it by myself. (Student 16, high-improvement learner, code group)

The other six students indicated that the feedback failed to meet their needs.
They wanted the feedback to provide examples and detailed explanations to illustrate errors and facilitate correction.

a. No, it does not. I mean... I want the feedback included more explanations and examples, so I know the problems and how to correct right away. But this one is not. (Student 10, low-improvement learner, code group)

b. I personally don’t like the feedback. I want the feedback to be as detail as possible. (Student 13, intermediate-improvement learner, code group)

c. Not good for me. It may be good for people who knows those errors well, but not for me. I wanted the feedback supports me to learn, not leave me there wondering what’s missing the sentence or how to correct it. (Student 15, intermediate-improvement learner, code group)

5.2.3. Triangulation of Quantitative and Qualitative Results

Data for the investigation of learner fit quality were drawn from the answers of 83 participants (41 from the corpus group and 42 from the code group) to the Likert-scale questionnaire and from the responses of 18 participants (9 from each group) to questions posed in semi-structured interviews. The quantitative results from the Mann-Whitney U tests with the questionnaire data showed that the corpus group had a significantly higher evaluation of corpus-based feedback on “clarity of feedback,” “appropriateness of feedback for task completion,” and “self-correction with the feedback” than the code group had when evaluating the coded feedback.

Students’ utterances about the learner fit of the two types of feedback from the semi-structured interview corroborated the quantitative results. Among the 28 utterance units from the corpus group, 89.29% were positive; there were 27 utterance units from the code group, of which only 29.63% were positive. Specifically, out of nine students, more than six participants in the corpus group felt that the feedback was easy to understand and was appropriate for them to complete the error correction practice exercises. Moreover,
eight students found that the corpus-based feedback met their personal needs and supported their individual learning paces and styles. In terms of the mixed and negative comments, two participants expressed a desire to access more examples and explanations in the corpus-based feedback. In the code group, two participants found the feedback fully understandable; three found it appropriate for the correction tasks and suitable to their individual needs. The remaining participants, however, voiced either mixed or negative comments. They felt the coded feedback was abstract and was unable to provide sufficient information to complete the error correction practice exercises. Moreover, their individual needs were not fully addressed by the feedback. Such introspective data from the interview, coupled with the questionnaire data, illustrate students’ perceptions of the two types of feedback on learner fit quality and the reasons behind their responses.

5.2. Impact of Corpus-based and Coded Feedback

After assessing the research results pertaining to learner fit for the two types of feedback, this study continued investigating the impact of the feedback on learning with the final research question:

- What evidence suggests that students treated with corpus-based feedback have a more positive learning experience than those treated with coded feedback?

Data addressing this question were collected from responses of the same 83 participants to twelve Likert-scale questionnaire items (41 participants from the corpus group and 42 from the code group) and from responses of the same 18 participants to questions posed in semi-structured interviews regarding their perceptions of the impact of learning with the target feedback. Analysis of the quantitative and qualitative data revealed that students perceived that corpus-based feedback exerted a more positive impact on their affective, cognitive, and intrinsic development towards error correction than the
coded feedback. However, the code group expressed interest in using other resources for self-learning after interacting with the coded feedback, an interest that was not voiced by the interviewees from the corpus group.

5.2.1. Questionnaire Data on Impact

The questionnaires assessed impact through twelve items designed to measure four sub-constructs: 1) satisfaction with the learning experience, 2) enthusiasm for using feedback to revise other errors, 3) error awareness, and 4) motivation for correcting syntactic errors in the future. Table 5.4 includes all twelve items and presents the descriptive statistics of the mean scores and the standard deviation of students’ responses to each item and each sub-construct.

Table 5.4 Questionnaire items on Learner Fit and participants’ responses to the Likert-scale items in the questionnaire

<table>
<thead>
<tr>
<th>No.</th>
<th>Questionnaire items</th>
<th>Corpus group N= 41</th>
<th>Code group N = 42</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1.</td>
<td>I am happy about my learning and improvement with the feedback.</td>
<td>6.07</td>
<td>.88</td>
</tr>
<tr>
<td>2.</td>
<td>After using the feedback, I am satisfied with my learning experience and improvement.</td>
<td>6.00</td>
<td>.84</td>
</tr>
<tr>
<td>3.</td>
<td>I am pleased with what I have learned from the feedback.</td>
<td>5.88</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>Affective: Enthusiasm for using feedback to correct other errors</td>
<td>6.26</td>
<td>.67</td>
</tr>
<tr>
<td>4.</td>
<td>I would be happy to receive the feedback on other grammatical errors for error correction.</td>
<td>6.29</td>
<td>.81</td>
</tr>
<tr>
<td>5.</td>
<td>I would welcome the feedback to help me correct other grammatical errors.</td>
<td>6.46</td>
<td>.78</td>
</tr>
<tr>
<td>6.</td>
<td>It would be nice to receive similar feedback on other grammatical errors to help me.</td>
<td>6.02</td>
<td>.72</td>
</tr>
</tbody>
</table>
Table 5.4 continued

<table>
<thead>
<tr>
<th>No.</th>
<th>Questionnaire items</th>
<th>Corpus group N= 41</th>
<th>Code group N = 42</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>7</td>
<td>The knowledge I have learned from the feedback will help me become aware of sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.80</td>
<td>1.10</td>
</tr>
<tr>
<td>8</td>
<td>After using the feedback in the error correction practice exercises, I am still unable to find sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.85</td>
<td>0.96</td>
</tr>
<tr>
<td>9</td>
<td>I have learned from the feedback about how to find sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.78</td>
<td>1.04</td>
</tr>
<tr>
<td>10</td>
<td>After using the feedback, I am motivated to correct sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.71</td>
<td>0.93</td>
</tr>
<tr>
<td>11</td>
<td>The learning experience with the feedback in the error correction practice exercises has given me the motivation to revise sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.76</td>
<td>0.92</td>
</tr>
<tr>
<td>12</td>
<td>The feedback I received has increased my interests in correcting sentence fragments, run-ons, and clause mistakes in my writing.</td>
<td>5.93</td>
<td>1.03</td>
</tr>
</tbody>
</table>

The Cronbach’s alpha of the twelve Likert-scale questionnaire items was .954, indicating a high level of internal consistency. The first and second sub-constructs—“satisfaction with the learning experience” and “enthusiasm for using feedback to correct other errors”—explored the affective impact of the feedback on language learning. The Cronbach’s alpha of the sub-constructs was .909 and .916 respectively. The third sub-construct was “error awareness,” referring to students’ perception of cognitive development achieved with the feedback. The scale also had a high level of internal
consistency, as demonstrated by a Cronbach’s alpha of .833. The last sub-construct was “motivation for error correction,” which evaluated students’ perception of their intrinsic learning motivation after interacting with the feedback. The Cronbach’s alpha of this sub-construct was .900.

The descriptive data for the total score of each sub-construct are summarized in Table 5.5.

Table 5.5 Total score and standard deviation of each sub-construct on impact

<table>
<thead>
<tr>
<th>No.</th>
<th>Questionnaire item</th>
<th>Corpus group N = 41</th>
<th>Code group N = 42</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>SD</td>
</tr>
<tr>
<td>1.</td>
<td>Affective: Satisfaction with the learning experience</td>
<td>17.95</td>
<td>2.33</td>
</tr>
<tr>
<td>2.</td>
<td>Affective: Enthusiasm for using feedback to correct</td>
<td>18.78</td>
<td>2.01</td>
</tr>
<tr>
<td></td>
<td>other errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Cognitive: Error awareness</td>
<td>17.44</td>
<td>2.54</td>
</tr>
<tr>
<td>4.</td>
<td>Intrinsic: Motivation for error correction</td>
<td>17.39</td>
<td>2.57</td>
</tr>
</tbody>
</table>

As seen in the table, the total scores of the items measuring the first and second sub-constructs were higher in the corpus group (Total $\text{corpus}_1$ = 17.95, SD $\text{corpus}_1$ = 2.33; Total $\text{corpus}_2$ = 15.26, SD $\text{corpus}_2$ = 3.21) than the code group (Total $\text{code}_1$ = 18.78, SD $\text{code}_1$ = 2.01; Total $\text{code}_2$ = 15.79, SD $\text{code}_2$ = 3.06). Similarly, the total scores of the items measuring the third and fourth sub-constructs were higher in the corpus group (Total $\text{corpus}_3$ = 17.44, SD $\text{corpus}_3$ = 2.54; Total $\text{corpus}_4$ = 14.60, SD $\text{corpus}_4$ = 3.18) than the code group (Total $\text{code}_3$ = 17.39, SD $\text{code}_3$ = 2.57; Total $\text{code}_4$ = 14.74, SD $\text{code}_4$ = 3.65).

A one-way ANOVA was conducted to evaluate whether students’ perceptions of impact for each type of corrective feedback were significantly different from each other. As
assessed by inspection of the boxplot in Figure 5.2, there were outliers among the total scores of each sub-construct of the corpus group.

![Boxplot of total scores of each construct concerning the impact of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21).](image)

Figure 5.2 Boxplot of total scores of each construct concerning the impact of each type of feedback in the 7-point Likert-scale questionnaire (Total Scores for each construct = 21).

The results gleaned from the original data with and without outliers were reported and compared. With the original data, the total scores of the corpus group’s satisfaction with learning (p = .000 < .05), enthusiasm for using corrective feedback for other errors (p = .000 < .05), error awareness development (p = .007 < .05), and motivation for error correction (p = .007 < .05) were not normally distributed when assessed by Shapiro-Wilk’s test. The total scores of the code group for satisfaction with learning (p = .257 > .05), enthusiasm for using corrective feedback for other errors (p = .174 > .05), and motivation for error correction (p = .057 > .05) were normally distributed, but those for error awareness were not (p = .035 < .05). Given that the assumption of homogeneity of variances for each construct were not met (P = .000 < .005 for each construct), as assessed by Levene’s
test of the homogeneity of variances, Welch’s F test was used to report the results. The results with one-way ANOVAs have shown that the scores of each sub-construct—satisfaction with learning (Welch’s F (1, 74.93) = 19.13, p < .0005, partial η2 = .18), enthusiasm for using corrective feedback for other errors (Welch’s F (1, 71.00) = 27.97, p < .0005, partial η2 = .25), error awareness development (Welch’s F (1, 77.99) = 20.32, p < .0005, partial η2 = .19), and motivation for error correction (Welch’s F (1, 73.70) = 14.72, p < .0005, partial η2 = .14)—were significantly different. The effect sizes indicated 18%, 25%, 19%, and 14% of variance in questionnaire scores of each sub-construct respectively can be attributed to feedback types.

When the outliers were deleted from the original data, there were no new outliers in the data set. The data in the new set on satisfaction with learning language (p = .025 < .05), enthusiasm for using corrective feedback for other errors (p = .005 < .05), error awareness development (p = .025 < .05), and motivation for error correction (p = .036 < .05) were not normally distributed, when assessed by Shapiro-Wilk’s test. The total scores of the code group on satisfaction with learning language (p = .257 > .05), enthusiasm for using corrective feedback for other errors (p = .174 > .05), and motivation for error correction (p = .057 > .05) were normally distributed, but those on error awareness were not (p = .035 < .05). The assumption of homogeneity of variances for each construct were still violated (p \text{ satisfaction} = .000; p \text{ enthusiasm} = .000; p \text{ error awareness} = .001; p \text{ motivation} = .001), as assessed by Levene’s test of the homogeneity of variances, and Welch’s F test was used to report the results. The results with one-way ANOVAs have shown that the scores for each sub-construct—satisfaction with learning (Welch’s F (1, 64.59) = 28.09, p < .0005, partial η2 = .25), enthusiasm for using corrective feedback for other errors (Welch’s F (1, 61.26) =
38.36, p < .0005, partial $\eta^2 = .31$), error awareness development (Welch’s F (1, 73.05) = 26.39, p < .0005, partial $\eta^2 = .23$), and motivation for error correction (Welch’s F (1, 67.52) = 19.95, p < .0005, partial $\eta^2 = .19$)—were significantly different. The effect sizes indicated 25%, 31%, 23%, and 19% of variance in questionnaire scores of each sub-construct respectively were attributed to feedback types. To sum up, the one-way ANOVAs using the data with and without outliers reached the same conclusion: that the corpus group had more positive perceptions about the affective, cognitive, and intrinsic impact of the feedback on learning. Students’ perceptions of the impact of the feedback on learning were further examined through their responses to the questions posed in semi-structured interviews.

5.2.2. Semi-structured Interview Data on Impact

The analysis of qualitative data revealed 82 idea units concerning four aspects of impact on learners, as seen in Table 5.6: learning satisfaction, learning interests, error awareness, and motivation.

Table 5.6 Positive and mixed/negative idea units identified with each theme and subtheme regarding impact, as reported in semi-structured interview transcripts

<table>
<thead>
<tr>
<th>Theme</th>
<th>Sub-theme</th>
<th>Corpus group N= 9</th>
<th>Code group N = 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Mixed/ Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>Satisfaction with learning experience</td>
<td>Satisfaction with learning experience</td>
<td>22.5% (9 units)</td>
<td>0 % (0 units)</td>
</tr>
<tr>
<td>Interest in learning about errors with the target feedback</td>
<td>Interest in learning about errors with the target feedback</td>
<td>20% (8 units)</td>
<td>2.5% (1 units)</td>
</tr>
</tbody>
</table>
The following sections discuss each theme concerning impact quality and substantiate the discussions with utterances from interview transcripts.

**5.3.2.1. Satisfaction with learning experience.** Students from both groups evaluated their satisfaction with their learning experience. All nine students from the corpus group expressed satisfaction.

a. Yes, it's good for me. I mean it's an interesting way to know those errors and correct them. (Student 1, low-improvement learner, corpus group)

b. Yes, very satisfied cause all my needs are met and I like the ways the system helped me learning. (Student 6, intermediate-improvement learner, corpus group)
c. Yes, generally good experience for me. (Student 9, high-improvement learner, corpus group)

Students in the code group, however, held varied opinions. Four reported being generally satisfied with the learning experience, as illustrated in the following statements.

a. Yes. I like this ways of learning grammars and errors. The exercises included different errors I usually made and the feedback taught me grammars. (Student 14, intermediate-improvement learner, code group)

b. Yes, I'm pretty satisfied cause I learned in that process. (Student 17, high-improvement learner, code group)

c. It's ok. I will say I can notice more errors than before. (Student 18, high-improvement learner, code group)

The other five, however, presented either mixed or negative opinions toward their learning experience with the coded feedback:

a. No, not at all. It's very hard to learn. (Student 11, low-improvement learner, code group)

b. To be frank, I'm not satisfied. I think the system can provide me more help, like more explanations and examples, but this one isn't. (Student 15, intermediate-improvement learner, code group)

c. Like 50% satisfaction. I mean I learned in the learning experience, but I sometimes get confused. (Student 13, intermediate-improvement learner, code group)

5.3.2.2. Interest in learning about errors with the target feedback. In addition to discussing their satisfaction/dissatisfaction with the learning experience, students discussed their interest in using target feedback to correct other errors in the future. Eight students from the corpus group believed that the corpus-based feedback would help them learn about other errors efficiently through examples, oral explanations, and colors. They also found learning with such feedback enjoyable. The following comments illustrate their enthusiasm for using the target feedback in the future.
a. Yes, I would like. If I can receive feedback like this in the future, I will learn more knowledge. (Student 2, low-improvement learner, corpus group)

b. Yes, like other sentence structure errors, prepositions, cause I want explanations like this. (Student 4, intermediate-improvement learner, corpus group)

c. Yes, it’s nice to learn with the feedback. (Student 8, high-improvement learner, corpus group)

Not all the students were unequivocally positive about the prospect of using corpus-based feedback to learn about other errors in the future. One student found it unnecessary to use corpus-based feedback for all types of error correction:

It depends. If the errors are simple, maybe not. If it’s hard, it will be helpful. [Researcher: Why? Can you explain more?] Like spelling, I do not need it, I can correct by myself, but some others difficult… like clauses, I feel it’s help. (Student 9, high-improvement learner, corpus group)

By comparison, four students in the code group believed the feedback would be useful for learning about other errors.

a. Yes, better than nothing. (Student 12, low-improvement learner, code group)

b. I think it’s helpful… [Researcher: Why?] Because it gives clues. (Student 14, intermediate-improvement learner, code group)

c. Yes, I want. If the feedback is not enough, I will check internet or textbooks, but it [the feedback] gives me directions. (Student 16, high-improvement learner, code group)

The remaining five students in the code group either disliked the feedback or voiced concerns with the idea of using the coded feedback on other errors:

a. I don’t really like it, to be honest. It doesn’t give explanations, so hard to understand. (Student 10, low-improvement learner, code group)

b. I’m not sure. It may work but may be confusing for some others. (Student 13, intermediate-improvement learner, code group)

c. Hard to say. If it is the only feedback, I will take it, but personally I don’t want to use it. [Researcher: Why?] Because I always want the feedback including more... and
could understand, but for me, it is hard. (Student 15, intermediate-improvement learner, code group)

5.3.2.3. Error awareness. In terms of whether and how the interaction with the feedback helped raise their awareness of errors, students’ comments were divided into two subcategories: awareness of syntactic errors in one’s academic writing and awareness of the importance of avoiding target syntactic errors in formal writing.

Under the first subcategory, six interviewees from the corpus group offered entirely positive comments on their increased error awareness, compared to four from the code group. Those six students from the corpus group felt that they were better at finding target syntactic errors in their own writing and that of their colleagues after practice with corrective feedback:

a. Yes, for example, I always say and write something that “I know how does it work.” I realize this is a big mistake and I have used it for a long time. (Student 2, low-improvement learner, corpus group)

b. Definitely. I know that the sentence fragments do not mean short sentence and run-on do not mean long sentence. There are some sentences I wrote are fragments and run-on sentences I used not to notice. (Student 4, intermediate-improvement learner, corpus group)

c. I think the better thing [after the error correction exercises] is to know errors by myself. I think I learned a lot and I find some those errors in my papers and in others. (Student 8, high-improvement learner, corpus group)

At the same time, three students from the corpus group still found difficulty with error awareness in their own writing despite using the feedback. They felt that they were unable to identify errors due to their own pre-existing writing habits.

a. I felt I noticed some, but not every time because it’s my writing and I always think it’s correct. (Student 1, low-improvement learner, corpus group)

b. If you tell me “this is a wrong sentence,” I can find it, but if not, I probably miss it. (Student 3, low-improvement learner, corpus group)
c. Not very easy for me to see those errors. It’s so hard to see those errors by myself. (Student 7, high-improvement learner, corpus group)

By comparison, in the code group, four students who held positive opinions remarked that the feedback had raised their awareness of syntactic errors in their own writing and that of others.

a. In general, I think I improved. I can see some errors without teachers’ comments. (Student 14, intermediate-improvement learner, code group)

b. I think so. I was unaware of my mistakes. For example, I used to use “there are” with a verb, but now I know it’s not. (Student 17, intermediate-improvement learner, code group)

c. What I found more interesting is that I found I can notice more errors in other people’s writing. Like last time, I read my classmate’s papers, I told him “you have many run-on sentences in your paper” and explained to him where they are. (Student 17, high-improvement learner, code group)

At the same time, five participants from the code group voiced mixed or negative responses about the feedback’s impact on error awareness. Three found that they did not learn from the feedback; as a result, the feedback was unable to exert positive impact on their writing.

a. No, it doesn’t [raise the awareness of errors in my writing]. The feedback is not good for me and I didn’t learn. (Student 10, low-improvement learner, code group)

b. I’m writing the way [I did before]. I don’t think anything changes. (Student 11, low-improvement learner, code group)

c. I kinda of aware of some run-on sentences, but it doesn’t really change others because I don’t understand frag and cl. (Student 13, intermediate-improvement learner, code group)

Two students said they were still unable to notice errors after the practice exercises due to the persistence of their own writing habits.

a. I could find some and correct them, but it’s not easy. If I can notice them, I will correct, but many times I couldn’t. [Researcher: Why?] Because I’m so used to it and
learning with the feedback can hardly change. (Student 15, intermediate-improvement learner, code group)

b. No. I could not find my own errors. (Student 16, high-improvement learner, code group)

In the second category, the awareness of the importance of avoiding target syntactic errors in formal writing, two interviewees from the corpus group and three from the code group noted the ways in which the feedback addressed the importance of avoiding the target errors in academic and formal writing:

a. And I used to write as what I say. But now I know it’s very important not to write run-on in my paper. I need to separate them or use words to make connections. (Student 5, intermediate-improvement learner, corpus group)

b. Another thing is I am now writing application letters for my internship. I guess I should make sure my sentences correct without clause errors and run-on sentences. (Student 8, high-improvement learner, corpus group)

c. I know that avoiding those errors are important. Like in speaking, fragments are okay, but in formal writing, we shouldn’t do that. Your readers will think you are not seriously at what you are writing. (Student 16, high-improvement learner, code group)

d. But, the exercises teaches me something, like to write correct sentences without errors, like fragments, run-ons. They are just not good. (Student 15, intermediate-improvement learner, code group)

5.3.2.4. Motivation for error correction and learning. During the interview, students also reflected on the ways in which the feedback and error correction exercises encouraged them to improve their writing accuracy and knowledge learning. The data analysis revealed that students’ utterances centered around three categories: motivation for self-correction, motivation to correct errors with teacher feedback, and motivation for self-learning.
Under the first category, motivation for self-correction, all participants in the corpus group reported increased motivation after using the target feedback.

a. Yes, the training motivated me. When I wrote a paper in 101C, I used to depend on teacher's feedback. After the training, I want to read my paper again and correct some run-on sentences and other errors by myself in my paper. (Student 8, high-improvement learner, code group)

b. I think so. I paid more attention to those [syntactic errors] now. I don't know how it's designed, but it includes many mistakes I often make. So when I work on my draft, I will not submit it after finishing the first draft. I feel that I should correct those errors first. (Student 5, intermediate-improvement learner, code group)

c. Yes, it's encouraging. (Student 9, high-improvement learner, corpus group)

In the code group, five students noted an increase in their own motivation after exposure to the feedback. They felt they were motivated by the learning experience to pay attention to and correct syntactic errors in their writing, as illustrated in the examples below.

a. Yes, I am. I feel it's my responsibility of taking care of those errors. (Student 12, low-improvement learner, code group)

b. Yes, the practice motivated me a lot. Usually, I was very lazy to read my sentences to check, but I recently I found some clause errors in my writing and correct them. (Student 13, intermediate-improvement learner, code group)

c. Yeah, I tried to correct them by myself. (Student 17, high-improvement learner, code group)

Two students in the code group found that the learning experience did not impact their motivation for error correction. They found themselves unmotivated due to their learning experience with the feedback and their own learning attitudes.

a. Maybe not very much, because I was not learning to correct them. (Student 11, low-improvement learner, code group)

b. It probably the same, not change really, probably I’m very lazy. (Student 10, low-improvement learner, code group)
Under the second subcategory, motivation to correct errors with teacher feedback, two students in the corpus group and two from the code group said that the feedback motivated them to deal with syntactic errors more seriously than before when receiving feedback from their instructor.

a. And when teacher highlighted my sentence as a fragment, I want to make sure I correct them correctly. [Researcher: Is what you are doing now different from your previous revision?] I previously don’t take it seriously. I just go and correct it. Now, because I learned, I want to make sure it’s correct. (Student 4, intermediate-improvement learner, corpus group)

b. After doing the exercises, I can better understand my teacher’s feedback on those errors and try best to correct them. (Student 3, low-improvement learner, corpus group)

c. Yeah, I corrected more errors when getting the feedback from my instructors. I know those errors should be treated seriously. (Student 12, low-improvement learner, code group)

d. Yes, my professor gave me the same feedback. So I understand it. (Student 14, intermediate-improvement learner, code group)

Under the third subcategory, motivation for learning, three students from the code group said that the feedback motivated them to learn more about syntactic errors:

a. It’s motivating. Actually, I don’t know how to correct clause errors in the exercises, so I learned more from internet. It has a lot of examples, and I understand the clause better. (Student 11, low-improvement learner, code group)

b. Because I felt that I didn’t learn well with the feedback, I tried to learn by myself. For example, I talked to my professor and asked her about sentence fragments, and clauses. (Student 15, intermediate-improvement learner, code group)

5.3. Triangulation of Quantitative and Qualitative Results

The triangulated quantitative and qualitative data clarified students’ attitudes towards the learning impact of both types of feedback. After Mann Whitney U tests were conducted, the quantitative analysis with data from the Likert-scale questionnaire revealed
that students in each group rated the impact of learning differently. Specifically, the mean scores for “satisfaction with the learning experience,” “enthusiasm for using feedback on other errors,” “error awareness raising,” and “motivation for error correction” from the corpus group were significantly higher than those from the code group.

The qualitative data on the impact of learning were collected from semi-structured interviews. In alignment with the results from the questionnaire analysis suggesting that students preferred corpus-based feedback to coded feedback, 87.10% of the utterances from the corpus group were positive, compared to 59.52% from the code group. Specifically, all nine interviewees in the corpus group found the learning experience with the corpus-based feedback to be supportive, efficient, and enjoyable; eight of them indicated that they were enthusiastic about using the feedback for learning and correcting other types of errors in the future, while one observed that it may be unnecessary to use corpus-based feedback for correcting errors such as spelling mistakes. By comparison, only four interviewees in the code group were satisfied with the learning experience and showed interest in using coded feedback for learning and correcting other errors. The remaining interviewees in the code group felt that the feedback did not provide sufficient information for learning; therefore, they were not enthusiastic about using such feedback for other errors. In terms of error awareness, six students from the corpus group found the feedback helpful in raising their awareness of errors in their own writing and that of their peers, while three still experienced difficulty maintaining awareness of errors on their own. In the code group, four students felt that the feedback had raised their awareness of errors in their own writing, but the other five students did not feel that way. In terms of learning motivation, all students in the corpus group found themselves motivated by the feedback.
Two of them also mentioned that they took teacher corrective feedback more seriously than before. By comparison, five students in the code group said that they were also motivated to identify errors by themselves, while two students noticed no such changes in themselves. Two students in the code group remarked that the feedback and training motivated them to correct syntactic errors when receiving teacher feedback. Three students from the corpus group found that they were motivated to learn knowledge related to syntactic errors by themselves: they either searched the internet or consulted their instructors about ways of correcting subordinate clause or fragment errors after the error correction exercises.

5.4 Chapter Summary

This chapter discussed findings related to two research questions corresponding to the learner fit and impact of corpus-based and coded feedback. In terms of learner fit, the quantitative analysis of Likert-scale questionnaire items indicated that students perceived the corpus-based feedback as being clearer, more generally useful, and more helpful for enabling self-correction than the coded feedback: The one-way ANOVA test indicated that the total scores for each of the three perception subconstructs (clarity of feedback, usefulness of feedback, and self-correction) was significantly higher for the corpus group than for the code group. Regarding the impact of learning, the one-way ANOVA test results also showed that students in the corpus group were more satisfied with the learning experience and more enthused at the prospect of using the target feedback to correct other errors. Moreover, the corpus group was more aware of the target syntactic errors and more motivated to correct sentence fragments, run-on sentences, and subordinate clause errors in their writing than the code group.
The semi-structured interviews provided further evidence concerning the learner fit and impact of the two types of corrective feedback. In terms of learner fit, 89.29% of utterances from the corpus group were positive. These comments indicated that the majority of students in the corpus group found the feedback easy to understand and appropriate for aiding task completion. Students also noted that the variety of features of the corpus-based feedback (e.g., examples, rule sections, colors, oral feedback) met their individual needs and facilitated individual learning. At the same time, 10.71% of the data consisted of mixed or negative comments. Those comments indicated that two students desired more examples and explanations. By comparison, only 29.63% of the utterances from the code group were positive. These commenters found the coded feedback fully understandable, appropriate for the correction tasks, and suitable to their individual needs. However, the remaining 70.37% of the utterances indicated that students preferred more detailed and informative feedback for error correction and task completion.

Different attitudes towards the impact of learning with each type of corrective feedback were also discernible in the interview. In total, 87.10% of comments from the corpus group from were positive, and 12.90% negative; 59.52% of comments from the code group were positive, while 40.48% were negative. All nine students from the corpus group said that they were pleased with their learning experience, and eight of them were eager to learn about other types of errors using the feedback. In the code group, only four students were satisfied with their learning experience and interested in learning about other errors with the coded feedback. In terms of error awareness, six students in the corpus group found that the feedback raised their awareness of errors in their own writing; two students also indicated that the feedback made them more aware of the importance of
reducing target syntactic errors in formal writing. However, three students still experienced difficulty noticing their own syntactic errors after the error correction practice exercises. In the code group, four participants remarked positively on the impact of the feedback for raising their awareness of errors, and three said they were more aware of the importance of avoiding target errors in academic writing after the exercises. However, five students did not find the feedback helpful for raising their error awareness: three attributed the unhelpfulness to their unsuccessful learning experience with the feedback, and two attributed it to their own ingrained writing habits. Regarding the impact on motivation for error correction and learning, nine students in the corpus group noted that the feedback and training motivated them to correct errors by themselves; two also said they were motivated to correct target errors with teacher feedback. In the code group, five students found themselves motivated for error correction, but two did not; two students also found that they took teacher feedback on syntactic errors more seriously. It is worth noting that two students from the code group also remarked that the coded feedback encouraged them to learn and explore syntactic errors out of class by themselves. Such responses from three semi-structured interview are consistent with the questionnaire data and provide in-depth insight into participants’ attitudes towards the learning impact of the two types of feedback.

The next chapter will synthesize answers to each research question and formulate connections among them. In addition, it will address the limitations of the study, its research implications, and recommendations for future work.
CHAPTER 6. CONCLUSION

This study described the development of a system of corpus-based feedback and evaluates its effects on students’ error correction and language learning. To evaluate the quality of corpus-based feedback, the study compared it to that of traditional coded feedback from three perspectives: language learning potential, learner fit, and impact. This chapter first synthesizes the key findings concerning these three aspects according to five research questions. Implications for the use of theoretical approaches for pedagogical design, the evaluation of CALL materials, and educational practices are then proposed, as are recommendations for future research. This chapter also reiterates the contextual information of the study to determine how best to interpret the findings and what the limitations of the study are. Based on the discussion, summative conclusions are drawn.

6.1. Summary of Findings Related to Each Research Question

Based on the logical framework for evaluating CALL corrective feedback and Chapelle’s evaluation criteria, five research questions have been developed concerning three qualities of CALL corrective feedback: language learning potential, learner fit, and impact. The key findings pertaining to each research question are summarized below.

6.1.1. Language Learning Potential

Language learning potential concerns the means by which the corpus-based feedback and the coded feedback have encouraged students’ error noticing, error correction, and language learning. Accordingly, three research questions were composed to explore these effects: 1) What evidence suggests that students with corpus-based feedback are better equipped to notice syntactic errors in the learning process than students with coded feedback? 2) What evidence suggests that students with corpus-based feedback are
better equipped to correct syntactic errors in the learning process than students with coded feedback? 3) What evidence suggests that students treated with corpus-based feedback are better equipped to apply syntactic knowledge to correct errors in a new text as an outcome of the learning than students treated with coded feedback? Each of these questions was addressed by analyzing quantitative and qualitative data. The quantitative data consisted of the scores from 90 students (45 from each test group) on error noticing and error correction in two exercises, as well as their pre- and post-test scores. The qualitative data consisted of responses to questions posed in semi-structured interviews with 18 participants (9 from each group). Overall, the analysis of both the quantitative and qualitative data indicated that corpus-based feedback was more beneficial than coded feedback for error noticing, error correction, and language learning.

The first aspect of language learning potential, error noticing with CALL corrective feedback, was evaluated through error noticing scores on two error correction exercises and students’ responses to questions in the think-aloud protocols and the semi-structured interview. Two one-way ANOVAs revealed that the error noticing scores of the corpus-based group in each exercise were significantly higher than those of the code group, indicating that the corpus group outperformed the code group in noticing syntactic errors. Moreover, the corpus group’s error noticing scores for sentence fragments, run-on sentences, and subordinate clause errors were higher than those of the code group. During the think-aloud protocols and semi-structured interviews, students from the two groups also voiced their attitudes towards the two types of corrective feedback on error noticing. In the corpus group, 96.48% of the comments remarked on the helpfulness of the corpus-based feedback for error noticing. Participants in the corpus group generally found the
colors, examples, and various features of the corpus-based feedback heightened their ability to notice syntactic errors in the exercises. By comparison, 58.43% of the utterances from the code group were positive. Half of the participants found the codes helped them notice errors in the sentences, while half still had difficulty finding target syntactic errors.

The second aspect of language learning potential, error correction with CALL corrective feedback, was evaluated through error correction scores on the error correction practice exercises and students’ answers in the think-aloud protocols and semi-structured interviews. The error correction scores of the corpus group were significantly higher than those of the code group. Specifically, the scores for each error type—sentence fragments, run-on sentences, and subordinate clause errors—in the corpus group were higher than those for each in the code group. Students’ utterances in the think-aloud protocols and semi-structured interview corroborated the quantitative results. Of the 90 remarks from the corpus group, 87.78% were positive, 12.22% negative. Students in the corpus group generally found that the feedback explained the errors clearly and was helpful for syntactic error corrections. The code group made 85 comments. Whereas 46.88% of these comments indicated that the coded feedback enhanced their error correction, 53.12% noted that the feedback was insufficient for error correction and generated confusion.

The learning gains with each type of feedback were also evaluated with both quantitative and qualitative data. The quantitative data was collected from the pre- and post-tests. Both groups had similar mean pre-test scores on correcting syntactic errors. Even though both groups had significant improvement on their syntactic knowledge from the pre-test to the post-test, the learning gains in the corpus group were statistically significantly higher than those in the code group. Both groups attributed their
improvement to the feedback in the error correction practice exercises. When discussing their unsuccessful learning experiences, both groups posited four main reasons for the lack of success: 1) failure to notice errors, 2) forgetting what had been learned through the feedback, 3) the length and complexity of sentence structures in the new article, and 4) not having learned with the feedback. Specifically, the primary reason for unsuccessful learning given by the corpus group was the failure to notice errors in the new text, while the main reason given by the code group was not having learned with the feedback.

6.1.2. Learner Fit

Learner fit concerns the appropriateness of the corpus-based and coded feedback for error correction. The research question pertaining to learner fit was: What evidence suggests that advanced-low ESL learners perceive that corpus-based feedback is more appropriate than coded feedback for enabling correction of syntactic errors? This question was answered by analyzing the responses of 83 students (41 from the corpus group and 42 from the code group) to nine Likert-scale questionnaire items and the responses of 18 participants (9 from each group) to questions posed in semi-structured interviews. The one-way ANOVA test showed that the mean scores of the corpus group on the three subconstructs of learner fit (clarity of feedback, appropriateness of feedback for task completion, and suitability of feedback for individual learner characteristics) were significantly higher than those of the code group.

Students’ responses during the semi-structured interview clarified their evaluation of the appropriateness of each type of feedback for correcting syntactic errors. Among the 28 utterance units from the corpus group, 89.29% were positive and 11.11% were negative. Participants in the corpus group generally found the feedback understandable
and appropriate for completing the task and meeting their individual needs. Among the 27 units from the code group, however, only 29.63% were positive, whereas 70.37% were negative. More than half of the interviewees from the code group found the feedback too abstract to understand, difficult to use to complete the task, and unsuitable for their individual needs.

6.1.3. Impact

The impact of learning refers to the influence of the corpus-based and coded feedback on students’ learning after the error correction practice exercises. The final research question targeted the impact quality of the two types of CALL corrective feedback on language learning: What evidence suggests that students treated with corpus-based feedback have a more positive learning experience than those treated with coded feedback? The responses of 83 participants (41 from the corpus group and 42 from the code group) to twelve Likert-scale questionnaire items and the responses of 18 participants (9 from each group) to questions posed in semi-structured interviews were analyzed to address this question. The data analysis indicated that the corpus group found that the feedback had a more positive affective, cognitive, and intrinsic impact on them than the code group. The one-way ANOVA test indicated that the mean scores of questionnaire items for “satisfaction with the learning experience,” “enthusiasm for using feedback on other errors,” “error awareness raising,” and “motivation for error correction” were significantly higher for the corpus group than the code group.

The 18 interviewees’ responses to the semi-structured interview offered complementary information to the quantitative data from the questionnaire. Participants in the corpus group uttered 31 idea units about the impact quality of the feedback. Of these,
87.10% (27 idea units) were positive, indicating that the students in the corpus group generally enjoyed learning about syntactic errors with corpus-based feedback and were willing to learn about other errors with it. In addition, more than half of the respondents in the corpus group noticed an improvement in their ability to identify target syntactic errors in their own writing and an increased motivation to correct these errors. However, three interviewees found that interacting with the feedback in the two exercises did little to change their writing habits or increase their error awareness. The responses from the code group contained 42 idea units concerning impact quality, of which 59.52% were positive. Specifically, more than half of the interviewees (five of the nine participants) were not fully satisfied with their learning experience, did not wish to use the feedback for correcting other errors, and did not become more aware of their own errors after the error correction practice exercises. Nonetheless, three interviewees mentioned that, although they failed to learn syntactic knowledge with the coded feedback, the feedback motivated them to explore target syntactic knowledge with online resources and their instructors.

6.2. Implications for Research, Educational Practice, Pedagogical Design, and CALL Material Evaluation

Corrective feedback has been widely used in language classrooms to help students identify and remedy grammatical errors. The effect of corrective feedback on students’ language learning, however, is still controversial. The findings and research design of this dissertation have extensive implications for corrective feedback research, language pedagogy, and the utility of theoretical approaches for tool development and learning evaluation.
6.2.1. Implications for Research

This dissertation exemplifies the use of quantitative and qualitative data in research evaluating the language learning potential of corpus-based feedback. The quantitative data reveal varying degrees of success for error correction and knowledge transfer using specific types of corrective feedback; the qualitative data provide detailed information concerning the reasons behind students’ successful and unsuccessful learning experiences. Despite a number of available studies on corrective feedback, studies using both quantitative and qualitative data to evaluate the efficiency of corrective feedback for language learning are sparse.

This dissertation contributes to the current research on corrective feedback by using both quantitative and qualitative data to explore the effects of corpus-based and coded feedback on error correction and language learning. Specifically, quantitative data were collected from error correction exercises and pre- and post-tests to evaluate immediate intake and language learning. The qualitative data from think-aloud protocols consisted of details surrounding students’ interactions with the feedback and features of the feedback that students perceived as facilitating error correction. Further data about learner perceptions, identified in the semi-structured interviews, addressed the reasons behind the success or failure of knowledge transfer. The triangulation of the quantitative and qualitative evidence not only provides insight into knowledge intake and integration, but also illustrates how students interacted with the target feedback and why corpus-based feedback worked more effectively explaining one error category than coded feedback.
The dissertation also has implications for exploring the effects of corrective feedback on correcting syntactic errors and teaching syntactic knowledge. Despite the large number of studies on corrective feedback, few have discussed its effects on learning about syntactic errors. Existing studies have found corrective feedback to be less helpful for correcting sentence-level errors (except for run-on sentences) than word-level errors (Ferris, 2006; Ferris & Robert, 2001; Sachs & Polio, 2007; Van Beuningen et al., 2008, 2012; Li & Hegelheimer, 2013). Truscott (2007) posited that corrective feedback offered little help with sentence-level errors due to their complexity. This dissertation focuses exclusively on the three most common syntactic errors in ESL students’ writing: sentence fragments, run-on sentences, and subordinate clause errors. The results of this study not only reflect the effectiveness of corpus-based feedback for noticing and correcting these errors, but also reveal that both the corpus and code groups had more success applying knowledge from the feedback to correcting run-on sentences than sentence fragments or subordinate clause errors. Such findings are in alignment with the results from one previous study (Li and Hegelheimer, 2013), which found that run-on sentences were easier to correct than sentence fragments. At the same time, the present study also explored the effects of corrective feedback on another type of syntactic error (i.e., subordinate clause errors) that was rarely discussed in previous research, and collected both quantitative and qualitative evidence to determine why students had the highest improvement correcting run-on sentences and the least improvement correcting subordinate clause errors. The extensive findings of this study include new evidence supporting the effectiveness of corrective feedback, and encourage further research into the effect of corrective feedback on students’ knowledge of syntactic errors.
6.2.2. Implications for the Utility of Theoretical Approaches for the Pedagogical Design of Corrective Feedback

The dissertation also has implications concerning the utility of theoretical approaches for the pedagogical design of corpus-based feedback. To date, many studies have discussed the development process of CALL programs and their corrective feedback design (e.g., Liou, 1991; Reuer, 2003; Milton, 2006; Milton & Cheng, 2010; Tokuda & Chen, 2004; Heift, 2004, 2010; Nagata, 1993; Shaalan, 2005; Yang & Akahori, 1998). Some studies describing the design of CALL programs (e.g., Yang & Akahori, 1998; Shaalan, 2005; Tokuda & Chen, 2004) did not articulate the principles or the theoretical foundations of the design of their feedback. In other studies (Heift, 2003, 2010; Milton, 2006; Milton & Cheng, 2010; Nagata, 1993; Reuer, 2003), researchers explained one or two aspects of the feedback design with linguistic theories. For instance, Reuer (2003) stated that the provision of tree structure error messages is informed by Lexical Functional Grammar (LFG) Theory. He noted that, since the tree structure can clearly identify errors and explain the syntactic properties of a sentence with similar terminologies used in traditional grammar teaching (e.g. subject, direct object), the feedback would be easily understood by users. Very few studies have explained their pedagogical design from SLA theoretical perspectives.

Recognizing the limitations of previous studies that theorize selected aspects of feedback design, the dissertation provides support for all aspects of its design and discusses the various features of the corpus-based feedback from different SLA theoretical perspectives. Specifically, the design of the corpus-based feedback was informed by two SLA theories: the Interactionist Approach and Skill-learning Theory. The Interactionist Approach, which emphasizes the negotiation of meaning in language learning, has
extended its discussion on the important role of negotiation of form for second language development. As a vehicle for the negotiation of form between L2 learners and the competent English speakers who designed the feedback, corpus-based feedback promotes negotiation of problematic language forms and the development of students’ error awareness and comprehension. Skill-learning Theory emphasizes the importance of introducing declarative knowledge to learners in language learning. Guided by this principle, corpus-based feedback uses written and oral metalinguistic feedback and rule sections to illustrate declarative knowledge to enhance learning. This study’s integration of the Interactionist Approach and Skill-Learning Theory lays a solid theoretical foundation for the design of corpus-based feedback and demonstrates the importance of utilizing theoretical approaches to maximize the pedagogical value of materials.

6.2.3. Implications for the Use of the Logic Model of the Theory of Action for Developing a Logical Framework to Evaluate CALL Materials

Discussing language program evaluation, Norris (2016) pointed out that “program evaluation enables a variety of evidence-based decisions and actions, from designing programs and implementing practices to judging effectiveness and improving outcomes” (p. 169). When conducting language program or CALL material evaluation, however, very few studies have used a framework to describe the interactive relationships among various actions that lead to certain learning outcomes and impacts (Norris, 2016).

Using a basic logic model of the Theory of Action, this study established a logical framework by articulating the relationship among theoretical foundations, CALL corrective feedback design, related activities for practice and learning, and key hypotheses about immediate intake, long-term learning outcomes, and learning impact. Applying Chapelle’s
criteria for CALL task appropriateness to corpus-based feedback evaluation, the study design not only addresses why and how each research question was composed, but also indicates the logical relationships among the questions. For example, three questions concerning language learning potential were composed to enable assessment of error noticing and error correction with CALL corrective feedback, and of knowledge improvement from the pre- to post-tests. The methods for assessing these three aspects of language learning were informed by the Interactionist Approach, which emphasizes the significance of learners’ noticing, comprehension of the input information, and knowledge integration. Guided by this theoretical approach, this study is one of only a few to have recorded all these aspects of learning when evaluating the effects of corrective feedback. It is recommended that future corrective feedback research establish its own evaluation framework and articulate the key elements of the framework, such as theoretical approaches and various hypotheses, which justify its CALL evaluation process and research question design.

6.2.4. Implications for Educational Practice

In addition to the implications for research, the development of corpus-based feedback and the My Feedback website has great potential to contribute to educational practice. First, the findings from this study underscore the possibilities of using My Feedback and corpus-based feedback to teach students about grammatical errors in academic writing courses. In this study, the corpus group learned ways of correcting target syntactic errors from the feedback provided on the error correction practice exercises and then applied their knowledge to a new text; interviewees in the corpus group were generally satisfied with their learning experience and outcomes, and they expressed their
enthusiasm for learning about other errors with corpus-based feedback. Such findings point to the need for further development of corpus-based feedback on a variety of error categories and corresponding error correction practice exercises to address various common errors in students’ writing. Moreover, when the website *My Feedback* is fully developed with the accompanying corpus-based feedback library, it will have the potential to improve the overall quality of corrective feedback and reduce teachers’ workloads. In a real classroom setting, the quality of written corrective feedback provided by instructors is limited by time constraints, class sizes, and high workloads. Once corpus-based feedback on a range of grammatical errors is developed, revised, and integrated into the *My Feedback* library, instructors will be able to extract existing corrective feedback from the website and insert it into students’ submissions as feedback on target errors. Instructors can also be trained to develop their own feedback banks and share them with other instructors through the feedback library. When grading students’ assignments on *My Feedback* websites, instructors stand to save a large amount of time normally spent composing feedback for certain types of errors, giving them more time to devote to other features of writing or course preparation.

### 6.3. Limitations and Suggestions for Future Research

Although the findings of the dissertation provide positive evidence for the use of corpus-based feedback for syntactic error correction, a number of factors and limitations restrict generalization of the results to other contexts. First, the participants of this study consisted exclusively of ESL students with English proficiency at the advanced-low level. These participants found that corpus-based feedback better facilitated their error noticing, error correction, and syntactic knowledge learning than coded feedback. In addition,
corpus-based feedback was found to be more appropriate for their learning and to inspire more positive affective, cognitive, and intrinsic motivation for language learning. The effectiveness of corpus-based feedback on students with different levels of English proficiency, however, should be tested and evaluated thoroughly rather than assuming the same results would be found for students in other contexts and at other ability levels.

Another limitation of the study relates to the use of think-aloud protocols for collecting qualitative data. During the think-aloud activities, students were asked to articulate what they were thinking when interacting with the feedback on the error correction exercises. Such introspective methods for collecting data have some inherent limitations that may be amplified when working with participants in their second language. The think-aloud data could reveal only what students were aware of while using the target feedback for error corrections and language learning. Their reports did not reflect any cognitive processes that may be important for learning even though they never reached students' consciousness. In the future study, the audio-data from the think-aloud protocols should be triangulated together with the video-data collected through think-aloud processes to gain insights into what may be unconscious during the time certain behaviors were engaged. Van Someren, Barnard, and Sandberg (1994) pointed out that differences exist when the participants of the think-aloud activities verbalize in their first or second language. The participants of the think-aloud protocols in this study were ESL learners. Even though they were advanced English language learners, they may have experienced difficulties verbalizing their thoughts and feelings about using target feedback for error corrections in English. Their utterances, therefore, may not precisely represent the students' thoughts as they would have expressed them in their first language. Moreover,
Norzayan, Nisbett, Smith, and Kim (2000) pointed out that cultural factors may hinder students’ thinking and talking processes in introspective research. They found students from East Asian countries were not used to sharing their ideas when they were involved in higher-order thinking. In this study, 12 out of 18 participants came from East Asian countries. To help mitigate this issue in the current research, a five-minute demonstration was performed by the primary researcher before the think-aloud activity to introduce think-aloud protocols and to illustrate how to verbalize one’s thoughts. Still, students found it challenging to describe their cognitive process aloud while simultaneously interacting with the feedback. Future research may include more extensive training sessions before the think-aloud activities to familiarize students with the tasks and promote verbalization skills.

The third limitation of the study is about its design. The study was designed with error correction practice exercises and pre- and post-tests. The error correction practice exercises were used to evaluate students’ ability to notice and correct errors immediately when interacting with target feedback; pre- and post-tests were used to assess students’ prior knowledge of syntactic errors, and their knowledge integration after treatment with either corpus-based or coded feedback. As Bitchener and Ferris (2012) have suggested, corrective feedback research may explore the ways in which error correction exercises with target feedback facilitate knowledge learning and retention over time. Due to time restraints, this study did not integrate delayed post-tests and other activities to explore the enduring effects of corrective feedback. Future studies on specific error categories might include delayed post-tests to provide further information on student knowledge retention over time. Studies might also include investigation into the effects of timing and frequency
of the corrective feedback for learning reinforcement. Furthermore, think-aloud protocols and semi-structured interviews could be integrated into future work to explore how students perceive their memory decline as time passes and how much detail students require from corrective feedback to be reminded of their own knowledge. Such research design will reveal the longitudinal effectiveness of corrective feedback and how many feedback treatments are needed for students to learn about specific types of errors.

Finally, while corpus-based feedback demonstrated advantages in facilitating knowledge intake and knowledge application to new texts in error correction exercises and pre- and post-tests, the effects of the feedback on students’ writing were not assessed in the dissertation. Granted, the study investigated students’ perception of the effects of corpus-based feedback on their own writing with Likert-scale questionnaires and semi-structured interviews. Students in the corpus group generally found that corpus-based feedback helped raise their error awareness of target syntactic error and strengthened their abilities to make corrections in their writing. However, during the semi-structured interviews, three students indicated that they still lacked confidence in their ability to detect errors in their own writing after the treatment due to the persistence of pre-existing writing habits. Considering this limitation, future studies can further explore the effects of corpus-based feedback on error correction in students’ writing. More specifically, samples of students’ writing before and after treatment with corpus-based feedback can be collected so that changes in the number of target syntactic errors can be observed.

6.4. Conclusion

Developing and evaluating corpus-based feedback is a time-consuming and expensive undertaking. The language learning potential, learner fit, and impact of corpus-
based feedback, however, are encouraging and exciting. Although both corpus-based and coded feedback improved students’ ability to correct the three types of syntactic errors in this study, the corpus-based feedback demonstrated greater strengths in promoting language learning, meeting individual needs, and exerting a positive impact on future learning. These findings have implications for future corrective feedback research.

This study suggests that researchers should collect both quantitative and qualitative evidence to evaluate the effects of corrective feedback on students’ performance and experience. This implication highlights the significance of this study to the corrective feedback literature, in which a large number of studies have only reported quantitative data to evaluate corrective feedback, but failed to include any qualitative evidence to address reasons for successful and unsuccessful error correction. Moreover, in contrast to the enduring claim that corrective feedback does not help users correct syntactic errors, the dissertation demonstrates the effectiveness of corpus-based feedback for correcting syntactic errors and facilitating syntactic knowledge learning. Given the scarcity of existing studies exploring the effectiveness of corrective feedback for correcting syntactic errors, the findings of this study are meaningful and encouraging, demonstrating the need for more studies exploring syntactic error correction with corrective feedback.

In addition to contributing to research on corrective feedback, this study sheds light on the design and evaluation of CALL materials. When explaining the development of CALL materials, very few studies have articulated guiding theories or specified their design principles. By contrast, this study exemplifies the use of different theoretical approaches for the pedagogical design of CALL materials. The use of SLA approaches from a cognitive
perspective promotes the quality of corpus-based feedback by adding features that facilitate students’ error correction and knowledge learning. The study findings suggest that designers of CALL materials ought to formulate their design using existing and proposed theoretical approaches rather than leave it ungrounded. Moreover, this study used the logic model of the Theory of Action for developing a framework for CALL material evaluation. Built on the logic model, the framework provides a means of visualizing the relationship among the theoretical foundations of the study, the pedagogical design of CALL materials, and all the learning processes that lead to the ultimate goal of learning. Therefore, this study contributes to CALL evaluation by using a logic model to justify its research questions and evidence collection. Despite some limitations, this study offers constructive guidance for future research and pedagogical applications.
REFERENCES


# APPENDIX A. ERROR AWARENESS SHEET

<table>
<thead>
<tr>
<th>ERROR CODE</th>
<th>TYPE OF ERROR</th>
<th>DEFINITIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>frag</td>
<td>Sentence Fragment</td>
<td>1) incomplete main clause; it may be a prepositional phrase or subordinate clause (sentence boundary issue)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2) Missing the verb in a clause</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3) Missing the subject in a clause</td>
</tr>
<tr>
<td>cl</td>
<td>Clause formation error</td>
<td>1) Improper coordination, subordination, or conjunction between clauses; or</td>
</tr>
<tr>
<td>cl-Ordin (Co- or Sub-)</td>
<td></td>
<td>2) incorrect word order; or</td>
</tr>
<tr>
<td>cl-adj</td>
<td></td>
<td>3) inappropriately omitted words or phrases, or unnecessary words or phrases in the clauses.</td>
</tr>
<tr>
<td>cl-adv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cl-noun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r-o/cs</td>
<td>Run-On Sentence</td>
<td>1) too many clauses or phrases without a logical break (sentence boundary issue)</td>
</tr>
<tr>
<td></td>
<td>Comma Splice</td>
<td>2) comma is used to join two independent clauses (sentence boundary issue)</td>
</tr>
</tbody>
</table>
APPENDIX B. PRE- AND POST-TESTS

First Name: ___________________________          Last Name: _______________________________

Error Correction Exercise: The original text includes various grammatical errors.

1. Please circle all the grammatical errors in the original text and make corrections in the
right column. (See Line 1)

2. If there is no error in a line, just leave it blank in the right column. (See Line 2)

<table>
<thead>
<tr>
<th>Original Text</th>
<th>Your Corrections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I know a Korean guy named Song-Lee from my English class at ISU. I am very</td>
<td></td>
</tr>
<tr>
<td>2 interested in Korean culture. Therefore, I</td>
<td></td>
</tr>
<tr>
<td>3 interviewed Song-Lee for this assignment.</td>
<td></td>
</tr>
<tr>
<td>4 South Korea is very far from the United</td>
<td></td>
</tr>
<tr>
<td>5 States. Even though, Song-Lee decided to</td>
<td></td>
</tr>
<tr>
<td>6 join a study abroad program and moved to</td>
<td></td>
</tr>
<tr>
<td>7 the U. S. when he was 16. When talking to</td>
<td></td>
</tr>
<tr>
<td>8 him about his native country, I have found</td>
<td></td>
</tr>
<tr>
<td>9 out a lot of interesting things from him.</td>
<td></td>
</tr>
<tr>
<td>10 Such as the patrilineal family structure in</td>
<td></td>
</tr>
<tr>
<td>11 Korea, Korean people’s loyalty to their</td>
<td></td>
</tr>
<tr>
<td>12 country, and their food culture.</td>
<td></td>
</tr>
<tr>
<td>13 First, Korean culture is patrilineal.</td>
<td></td>
</tr>
<tr>
<td>14 Many Korean people prefer to have a son</td>
<td></td>
</tr>
<tr>
<td>15 over a daughter. They believe that if their</td>
<td></td>
</tr>
</tbody>
</table>

Know
son gets married and has a child, then the family name gets passed to the next generation. While a daughter cannot pass on the family name. Take Song-Lee as an example, he was born in a traditional Korean family and has two elder sisters. As the only son in the family, Song-Lee gets great attention and support from his parents, however, he is also expected to take care of all of his family members in the future. All of those pressures and expectations are placed on him by his parents and Korean society, but they have also become motivation for him to study abroad to gain knowledge and experience. As Song-Lee told me that Koreans are patriotic. Song-Lee was born in Gwang Ju, South Korea. His hometown holds a lot of social events that influence him to this day, even he is far from home. For instance, The March 1st holiday, which is held to commemorate the Korean Sam-il
Movement. It is a famous Korean holiday. The Korean Sam-il Movement began on March 1, 1919 and lasted 12 months. Thousands of Koreans against Japanese domination and participated in more than 1500 demonstrations. Today, March 1st is a national holiday in South Korea. In Gwang Ju, there are lots of people celebrate this holiday. On this day, people hang the Korean national flag in front of their houses, that shows citizens' loyalty to the nation. Also, all of the family members gather together and pray for the ancient Korean people died for their country's freedom in the movement. The elders tell stories about how did the Korean soldiers protect their country in the past. Although those soldiers and heroes died long ago, their sacrifice will never be forgotten, they will live on in people’s hearts. Song-Lee said he is very patriotic and is influenced by those social events. He wants to go back
to South Korea immediately after graduation. Because if he stays in the U.S., he thinks that he cannot serve his country well.

In addition, Korean food is very healthy and nutritious. For example, kimchi is a traditional Korean side dish which is made of a variety of Korean seasonings and vegetables. In the article “World’s Healthiest Foods: Kimchi”, by Joan Raymond, explains that “kimchi is loaded with vitamins A, B, and C, but its biggest benefit may be in its ‘healthy bacteria’”. Especially people feel bloated after eating, they can consume a little kimchi and the good bacteria will help with digestion. Additionally, kimchi can help people reduce the risk of obesity, it can also lower cholesterol levels in people’s bodies. Korean people consume an enormous amount of kimchi, they eat kimchi as a side dish at every single meal,
<table>
<thead>
<tr>
<th>Line</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>83</td>
<td>they even use kimchi to make kimchi rice.</td>
</tr>
<tr>
<td>84</td>
<td>Song-Lee loves cooking dishes with kimchi,</td>
</tr>
<tr>
<td>85</td>
<td>when he stays in his apartment, he makes</td>
</tr>
<tr>
<td>86</td>
<td>kimchi soups or noodles and invites his</td>
</tr>
<tr>
<td>87</td>
<td>friends to have a meal together.</td>
</tr>
<tr>
<td>88</td>
<td>In conclusion, I have learned many</td>
</tr>
<tr>
<td>89</td>
<td>different aspects of South Korean culture</td>
</tr>
<tr>
<td>90</td>
<td>from Song-Lee. Korean families are mainly</td>
</tr>
<tr>
<td>91</td>
<td>patrilineal, Korean people have their own</td>
</tr>
<tr>
<td>92</td>
<td>ways of showing devotion to their country</td>
</tr>
<tr>
<td>93</td>
<td>they also have an interesting food culture.</td>
</tr>
<tr>
<td>94</td>
<td>By interviewing Song-Lee allows me to</td>
</tr>
<tr>
<td>95</td>
<td>learn about Korean cultures and history.</td>
</tr>
</tbody>
</table>
APPENDIX C. BACKGROUND SURVEY

First name: _____________________    Last name:__________________

1. How old are you?

____________________

2. What is your gender?

☐ Female
☐ Male

3. What is your first language?

☐ Arabic
☐ Chinese
☐ Korean
☐ Malay
☐ Spanish
☐ Turkish
☐ Vietnamese
Other (Please specify: ________________)

4. What is your level of study?

☐ Undergraduate
☐ Graduate
Other (Please specify: ________________)

5. Which college do you belong to?

☐ Engineering (Aerospace, Chemical and Biological Engineering, Civil, Construction and Environmental Engineering, etc.)
☐ Business
☐ Liberal Arts and Science (Political Science, English, Anthropology, etc.)
☐ Agriculture and Life Science (Agriculture, Animal Science, Economics, Statistics, etc.)
☐ Human Science (Apparel, Family and Consumer Science, Diet and Exercise, Kinesiology, Hospitality, etc.)
☐ Design (Architecture, Art & Design, Landscape, etc.)
☐ Veterinary Medicine
☐ Undecided/Not sure

6. What was your score on the TOEFL or IELTS (If you haven’t taken these tests, please leave the answer blank.)

TOEFL IBT Total score: ____________ Writing section: ______________

TOEFL PBT Total score: ______________

IELTS Total score: ______________ Writing section: ______________

7. How long have you been studying in the United States?

☐ Less than 6 months
☐ Between 6 months and 1 year.
☐ Between 1 and 2 years.
☐ Between 2 and 4 years
☐ More than 4 years
APPENDIX D. QUESTIONNAIRE

Please answer the following questions about your experience of and feelings about the use of corrective feedback for error corrections and language learning. Please choose answers truthfully and completely to the best of your knowledge.

1. The feedback helped to explain the errors clearly.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

2. After using the feedback in the error correction practices, I am still unable to find sentence fragments, run-ons, and clause mistakes.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

3. The learning experience with the feedback in the error correction practices has given me the motivation to revise sentence fragments, run-ons and clause mistakes.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

4. I found the feedback was useless in helping me correct sentence fragments, run-ons and clause mistakes in the error correction practices.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

5. The feedback was understandable in the error correction practices.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

6. The feedback I received has increased my interests in correcting sentence fragments, run-ons and clause mistakes.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

7. The feedback was useful for teaching me how to correct the mistakes in the error correction practices.
   Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree
8. The feedback was helpful for correcting sentence fragments, run-ons and clause mistakes in the error correction practices.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

9. The feedback helped me to revise the error correction practices by myself.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

10. I had difficulties in understanding the feedback.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly agree

11. I could use the feedback to correct sentence fragments, run-ons and clause mistakes in the error correction practices without any additional help.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

12. The knowledge I have learned from the feedback will help me become aware of sentence fragments, run-ons, and clause mistakes in my writing.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

13. I am happy about my learning and improvement with the feedback.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

14. I have learned from the feedback about how to find sentence fragments, run-ons, and clause mistakes.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

15. The feedback itself was enough to help me revise mistakes in the error correction practices on my own.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

16. After using the feedback, I am satisfied with my learning experience and improvement.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree
17. I would welcome the feedback to help me correct different grammatical errors in my writing.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

18. I am pleased with what I have learned from the feedback.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

19. It would be nice to receive the feedback to help me with my writing.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

20. I would be happy to receive feedback on various grammatical errors in the future.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree

21. After using the feedback, I am motivated to correct sentence fragments, run-ons and clause mistakes.

Strongly Disagree ①------- ②------- ③------- ④------- ⑤------- ⑥------- ⑦ Strongly Agree
APPENDIX E. THINK_ALOUD PROTOCOL INSTRUCTIONS

You are going to use the feedback in the system to correct errors in an error correction exercise.

1. Please say whatever you are looking at, whatever you are thinking, and whatever you are feeling when interacting with the feedback.

2. You can stop the task at any time when you feel uncomfortable.

3. This activity does not intend to investigate your performance. We are evaluating the feedback you receive. Any difficulties that you might be experiencing are not your fault.

4. Please be aware that the session will be audio-recorded so that researchers can go back and refer to what participants did and how they reacted.
APPENDIX F. IRB APPROVAL

The project referenced above has received approval from the Institutional Review Board (IRB) at Iowa State University according to the dates shown above. Please refer to the IRB ID number shown above in all correspondence regarding this study.

To ensure compliance with federal regulations (45 CFR 46 & 21 CFR 56), please be sure to:

- Use only the approved study materials in your research, including the recruitment materials and informed consent documents that have the IRB approval stamp.
- Retain signed informed consent documents for 3 years after the close of the study, when documented consent is required.
- Obtain IRB approval prior to implementing any changes to the study by submitting a Modification Form for Non-Exempt Research or Amendment for Personnel Changes form, as necessary.
- Immediately inform the IRB of (1) all serious and/or unexpected adverse experiences involving risks to subjects or others; and (2) any other unanticipated problems involving risks to subjects or others.
- Stop all research activity if IRB approval lapses, unless continuation is necessary to prevent harm to research participants. Research activity can resume once IRB approval is reestablished.
- Complete a new continuing review form at least three to four weeks prior to the date for continuing review as noted above to provide sufficient time for the IRB to review and approve continuation of the study. We will send a courtesy reminder as this date approaches.

Please be aware that IRB approval means that you have met the requirements of federal regulations and ISU policies governing human subjects research. Approval from other entities may also be needed. For example, access to data from private records (e.g., student, medical, or employment records, etc.) that are protected by FERPA, HIPAA, or other confidentiality policies requires permission from the holders of those records. Similarly, for research conducted in institutions other than ISU (e.g., schools, other colleges or universities, medical facilities, companies, etc.), investigators must obtain permission from the institution(s) as required by their policies. IRB approval in no way implies or guarantees that permission from these other entities will be granted.

Upon completion of the project, please submit a Project Closure Form to the Office for Responsible Research, 202 Kingland, to officially close the project.

Please don't hesitate to contact us if you have questions or concerns at 515-294-4566 or IRB@iastate.edu.