PARALLEL CORPORa AND PEDAGOGY:
ENHANCING CHINESE FOREIGN LANGUAGE LEARNING EXPERIENCE
THROUGH PARALLEL CORPUS TECHNOLOGY

A Dissertation in
Applied Linguistics and Asian Studies
by
Brody T. Bluemel

© 2015 Brody T. Bluemel

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2015
The dissertation of Brody T. Bluemel was reviewed and approved* by the following:

Xiaofei Lu  
Gil Watz Early Career Professor in Language and Linguistics and Associate Professor of Applied Linguistics and Asian Studies  
Dissertation Advisor  
Co-Chair of Committee  
Chair of the Graduate Program

Ning Yu  
Professor of Applied Linguistics and Asian Studies  
Co-Chair of Committee

Karen E. Johnson  
Kirby Professor in Language Learning and Applied Linguistics

Robert W. Schrauf  
Professor of Applied Linguistics

* Signatures are on file in the Graduate School.
ABSTRACT

The past several years have witnessed the continued advancement of corpora in linguistic research and language pedagogy. More recently, scholars have begun to explore the application of specialized corpora, such as parallel corpora. The application of parallel corpora in foreign language pedagogy remains relatively underexplored despite their apparent potential. The unique bilingual aligned structure of parallel corpus texts suggests their specific value in aiding students in writing and reading comprehension. These two aspects of language development are also key struggles for many second language learners of Chinese, as the complex nature of Chinese orthography makes the task of written language acquisition especially challenging. The current study capitalizes on the unique features of parallel corpora in the creation of a new parallel corpus language-learning tool to assist learners in the development of Chinese language writing and reading comprehension. This parallel corpus presents learners with texts in four formats: Chinese characters, Chinese characters + tone marks, pinyin, and English. The study then specifically addresses three primary issues: 1) if and how the parallel corpus aids learner development, 2) how both the students and instructor implement it, and 3) student opinion of the corpus tool. The application of the corpus takes place in two beginning-level high school Chinese classrooms that were both exposed to the same treatment. The study takes place over a three-month period, and analyzes participant assignments, exams, learning logs, a final questionnaire, screen-recordings of corpus use, and an instructor autoethnography through the Sociocultural Theoretical (SCT) framework to observe the affects of the implementation of the parallel corpus. An overall positive learning outcome is observed, with participants experiencing
marked development when approaching learning tasks conceptually and using the parallel corpus as directed. The study not only demonstrates how a parallel corpus can promote the linguistic and conceptual development of learners, but also provides insight into how such technology can more effectively be incorporated and adapted in foreign language pedagogy. The results further inform the methodological practices of parallel corpus development and pedagogy by providing an innovative new approach to corpus design, application, and theoretical analysis.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................ vii
LIST OF FIGURES ...................................................................................................................................... viii
ACKNOWLEDGEMENTS ................................................................................................................................. ix

## CHAPTER 1 – INTRODUCTION ..................................................................................................................... 1
  1.1. Parallel Corpora and Pedagogy ........................................................................................................ 1
  1.2. Sociocultural Theoretical Analysis of Development ................................................................... 4
  1.3. Purpose of the Study ....................................................................................................................... 6
  1.4. Research Questions ......................................................................................................................... 8
  1.5. Organization .................................................................................................................................... 10

## CHAPTER 2 – LITERATURE REVIEW ........................................................................................................ 12
  2.1. Parallel Corpora in Linguistic Research ....................................................................................... 13
  2.2. Parallel Corpora in Language Learning ........................................................................................ 18
  2.3. Pedagogical Applications of Parallel Corpora ............................................................................ 23
  2.4. Sociocultural Theory ..................................................................................................................... 34

## CHAPTER 3 – RESEARCH DESIGN AND METHOD .................................................................................. 41
  3.1. Context of the Study ....................................................................................................................... 43
    3.1.1. Setting .................................................................................................................................... 43
    3.1.2. Participants ............................................................................................................................. 44
    3.1.3. Course Details ....................................................................................................................... 45
  3.2. The Parallel Corpus Tool ............................................................................................................... 48
    3.2.1. Interface and Design .............................................................................................................. 48
    3.2.2. Corpus Materials .................................................................................................................. 53
  3.3. Data Analysis .................................................................................................................................. 56
    3.3.1. Data ....................................................................................................................................... 56
      3.3.1.1. Screen Recordings ........................................................................................................... 57
      3.3.1.2. Learner Logs and Questionnaire .................................................................................... 58
      3.3.1.3. Instructor Autoethnography .......................................................................................... 59
      3.3.1.4. Course Material .............................................................................................................. 60
    3.3.2. Analytical Framework/ Approach ......................................................................................... 63

## CHAPTER 4 – DATA ANALYSIS ............................................................................................................... 67
  4.1. Course Material and Screen Recordings Analysis ....................................................................... 69
    4.1.1. Poor Performers ..................................................................................................................... 70
      4.1.1.1. Month 1 ........................................................................................................................... 71
      4.1.1.2. Month 2 ........................................................................................................................... 73
      4.1.1.3. Month 3 ........................................................................................................................... 76
    4.1.2. Resistant to Parallel Corpus ..................................................................................................... 78
      4.1.2.1. Month 1 ........................................................................................................................... 79
      4.1.2.2. Month 2 ........................................................................................................................... 83
LIST OF TABLES

Table 2.1. Tokens of skal ................................................................. 16
Table 2.2. Concordance Search Results ............................................ 26
Table 3.1. Course Schedule .................................................................. 47
Table 3.2. Tone Marks in Mandarin Chinese ........................................... 50
Table 3.3. Data Type and Description .................................................... 56
Table 3.4. Course Material Data ........................................................... 61
Table 4.1. Categories According to Participant Performance ...................... 69
Table 4.2. Participant Percentage Grades ............................................... 70
Table 4.3. Ratings of Parallel Corpus ...................................................... 131
LIST OF FIGURES

Figure 3.1. The Four Text Formats ........................................................................................................ 51
Figure 3.2. Parallel Corpus Layout .......................................................................................................... 55
Figure 4.1. Example of Vocabulary Section of Exam .............................................................................. 72
Figure 4.2. Adam’s Parallel Corpus Pinyin Search ................................................................................. 74
Figure 4.3. Jack’s Parallel Corpus Search Using iPad Drawing Input .................................................... 77
Figure 4.4. Cathy’s Use of Interactive Character + Tone Mark Feature ...................................................... 84
Figure 4.5. Cathy’s Use of Interactive Pinyin Feature ............................................................................... 84
Figure 4.6. Cathy’s Search in Chinese Using nciku .................................................................................. 85
Figure 4.7. Kate’s Use of the Parallel Corpus Using and iPad ................................................................. 87
Figure 4.8. Kate’s Query using Chinese Input ............................................................................................ 88
Figure 4.9. Kate Interacting with Returned Search Results ....................................................................... 88
Figure 4.10. Isabel Searching Lesson 15 in the Parallel Corpus During the Exam .............................. 89
Figure 4.11. Gail’s Unsuccessful Parallel Corpus Pinyin Search ................................................................. 102
Figure 4.12. Ellen’s Use of the Parallel Corpus During the Zodiac Assignment ...................................... 103
Figure 4.13. Ellen’s Use of the Parallel Corpus During Exam 2 ................................................................. 104
Figure 4.14. Leo’s Search Using the Parallel Corpus During the Exam ..................................................... 108
Figure 4.15. Ellen’s Use of the Parallel Corpus During Exam 3 ................................................................. 109
Figure 4.16. Ellen’s Final Presentation Script ............................................................................................ 111
Figure 4.17. Gail’s Final Presentation Script ............................................................................................ 111
Pursuing a Ph.D. and completing a dissertation has been a very challenging and yet rewarding process, and I am truly grateful for all those who have helped me along this journey. I have learned from and felt great support from faculty, mentors, colleagues, students, family, and friends. I am indebted to each and every individual who has made it possible for me to learn and grow as an individual during this process and to be able to accomplish this goal in my education, and set the path for my future career.

I would like to especially acknowledge the individual support and guidance from my advisor and doctoral committee. I thank Dr. Xiaofei Lu, my advisor, for his continuous guidance during the entire process, and his inspirational example. I was always amazed at his ability to provide the exact support I needed without ever complicating the process, and especially at his immediate response to any question or problem that arose. His professional example is something I hope to continue to learn from and to emulate. I thank Dr. Karen Johnson for mentoring me, for always being willing to meet and discuss my research project, and for having a seemingly innate ability to make complex issues become clear. Dr. Ning Yu helped me develop professionally, and from the moment he joined the faculty he made every effort to provide me with further scholarly and professional opportunities that both enabled and enhanced my research and professional development. I am also grateful to Dr. Robert Schrauf for advising me during the beginning years of my study, and then continuing to assist me professionally both in my assistantship responsibilities and by opening paths for me to complete my research.
Additionally, I would like to thank my colleagues and friends who supported me throughout this process. In particular, Matthew Jadlocki and Dorothy Worden were great friends and intelligent minds that provided me with many great suggestions, ideas, and reassurances that helped improve my research project, and gave me confidence in the pursuit of my endeavors.

I am also indebted to Dr. Wen-hua Du, Chunyuan Di, and Shirley Donovan for the roles they played in providing me with the opportunity to work with the local high school to develop the Chinese program, and ultimately complete my research there. I also thank the many students who I had the pleasure of both teaching and learning from. While my intention was to teach them and guide them in their development, I feel as though I gained much more from the experience. Each individual student helped me to develop as a teacher, and to learn countless lessons in reviewing the data from their participation in this research study. I look forward to hearing of their accomplishments as they now enroll in college and begin figuring out their futures.

My wife Brittany has undoubtedly provided me with the greatest support through this entire process, and experienced the ups and downs together with me along the way. I am truly grateful for her support of both me and our children, and enabling me to pursue and now accomplish my educational goals. Our two children, Anders and Corlista, who joined our family during this process, have also been a continued inspiration to me. I am grateful for their love and support, and for my other family and friends who have also sustained me along the way.
CHAPTER 1
INTRODUCTION

The premise of this project grew out of the desire to address the challenge of written language acquisition amongst learners of Chinese. As both a student and instructor of Chinese, the disconnect between character acquisition and reading and writing comprehension has always been very transparent. Whereas alphabetical languages afford learners the ability to sound out words and draw on aural recognition for terms they have heard but not yet seen, Chinese necessitates the identification of all characters in a text before an individual can access the meaning or pronunciation. This disconnect leads to delayed reading and writing development and frustration as learners must have a substantial lexicon before they can read through even basic texts. This challenge of character acquisition prompted the researcher to consider several approaches to conceptual development of written Chinese, before identifying parallel corpora as a promising resource to explore and adapt in the Chinese foreign language classroom.

1.1. Parallel Corpora and Pedagogy

Parallel corpora allow users to conceptualize writings in the target language through the established schemas of their first language (L1), by providing aligned parallel representations of text in two (or more) languages. Parallel corpora are defined as specialized translation corpora that consist of source texts in one language that are aligned with translation texts in one or more additional languages, and they belong to the subcategory of language corpora known as translation corpora. As McEnery and Xiao (2008) explain, translation corpus is an umbrella term that includes parallel corpora, comparable corpora and equivalent corpora. While parallel corpora are constructed of
aligned source texts and translations in two (bilingual corpora) or more languages (multilingual corpora), comparable corpora are composed of two or more comparable monolingual sub-corpora that are designed using the same sampling frame, and equivalent corpora are a combination of these first two categories.

Parallel corpora are constructed using either “unidirectional, bidirectional, or multidirectional” source texts (McEnery & Xiao 2008, p. 20). The unidirectional design includes original source texts from one language, and translations in the other (e.g. English source texts/ Chinese translations or Chinese source texts/ English translations). Bidirectional parallel corpora include a balance of source texts from both languages and their translations (e.g. both English source texts/ Chinese translations and Chinese source/ English translations). Multidirectional parallel corpora are multilingual corpora that include the same piece of writing in numerous languages (e.g. the same article in Chinese, English, and German). These source texts and translations are then aligned in order to link the corresponding links together. Alignment can be accomplished using phrasal alignment, sentential alignment, or context-based alignment (Biçici 2008).

Phrasal alignment links set phrases in the corresponding languages together (often at the word level). Sentential alignment links the two translations together in sentence segments. Context-based alignment requires several steps of machine processing in order to determine contextual phrases in the two languages, and then links the corpora together based upon the concept or idea being expressed (Biçici 2008, p. 435). Consideration of alignment method becomes important in considering how parallel corpora are used, as the different levels of alignment provide a different interaction with the texts for the parallel corpus user or language learner.
Conventionally, parallel corpora have primarily been used for research in comparative linguistics and translation studies (Aijmer & Altenberg 1996; Aston 1997; Johansson 1999). Exploration of the pedagogical application of this technology is a more recent trend, with a developing field of research. The promising results of the pedagogical research that has been published, though, have led scholars to call for both better developed parallel corpora and additional pedagogical studies (Barlow 2000; Johansson 2009; Wang 2001; Fan & Xu 2002). The current study responds to this call by both developing a new web-based Chinese/English parallel corpus tool designed specifically for pedagogical application (www.parallelcorpus.com), and then researching its application in a Chinese foreign language classroom. This corpus tool was designed with the intent of capitalizing on the unique features of parallel corpora in order to more effectively assist language learners in the development of reading, writing, and lexical acquisition in their target language. As mentioned, Chinese has a uniquely complex orthography that is a recognized challenge for both native and foreign language learners, and the features of www.parallelcorpus.com seek to make this complex writing system more accessible to learners through the use of the Romanized pinyin system, and the use of the learners’ L1.

The parallel corpus tool created for this study was designed for both research and pedagogical application, but with a primary focus on creating a resource that provides more practical and fluid methods for effective and efficient written language acquisition. The typical format of parallel corpora enables the user to search a word in either language (in this case either Chinese or English), and the concordance query results display all sentences that contain that word – as well as all of the corresponding translations in the
second language. This approach to language analysis and study allows the researcher and/or learner to see how a word or phrase is used in a target language, and also to identify how that same expression is made in the corresponding or ‘parallel’ language. The corpus created for this project offers additional features beyond the typical platform, by allowing users to view the aligned texts not just in Chinese and English, but also in two additional formats: Chinese characters with tone marks and Romanized pinyin. This unique approach makes writings more accessible to language learners by providing texts in a format they are able to read (as the pinyin format aids them in identifying characters unknown to them).

The pedagogical application of this corpus took place in two beginning-level high school Chinese classrooms over a three-month period. During the first month of the study participants learning experience remained unaltered. The parallel corpus was not introduced until the second month of the study, at which time students were instructed on its use and applications and encouraged to apply it. During the final month of the study use of the parallel corpus was required in order to complete several tasks. The students had access to the corpus tool both in-class and outside of class to study and to complete a series of assignments and learning activities. The study followed their progression and analyzed performance and reported experience over this three-month period.

1.2. Sociocultural Theoretical Analysis of Development

The study approached analysis from a Sociocultural Theoretical (SCT) perspective, using the theory to describe if and how the parallel corpus tool affected learner development. SCT describes development as a dynamic social process that is mediated by culture, context, language, and society, among other possible artifacts or
learning tools (Lantolf & Thorne 2006; Johnson 2009). Furthermore, development is evidenced by the internalization of concepts, or conceptual development, as demonstrated through an individuals’ ability to appropriately apply the acquired concept or knowledge in new and different settings. Accordingly, the study sought to not only identify instances where the parallel corpus aided learner experience, but to observe how the acquisition of new concepts through the parallel corpus transformed the learner and their approach to Chinese language study. SCT recognizes that optimal development takes place when a learner functions within their Zone of Proximal Development (ZPD), which is understood as the distance between what an individual can achieve alone versus with mediation (Lantolf 2000b). The adaptation of a mediational tool, such as the parallel corpus, should therefore aim to promote development by providing a learner with the necessary mediation to allow them to function at the top of their ZPD, ultimately resulting in greater development and achievement. In order to make greater strides in learning and development, an individual must also be consistently challenged and presented with cognitive struggles that must be navigated to develop understanding. One approach taken within the SCT framework to promote such learning is Concept-Based Instruction (CBI). CBI emphasizes top-down transfer, in which a learner is presented with a complete concept and then provided with the necessary mediation to fully understand and later use that concept. This theoretical framework provides great insight into how the parallel corpus affects development, as the layout of parallel corpora presents learners with complete textual samples of language use in the target language aligned with those same concepts represented in their L1, allowing them to use their L1 to mediate their understanding of the target language.
1.3. Purpose of the Study

The overall purpose of the study was to identify if, and to what extent, the parallel corpus aided students in their acquisition and interaction with written Chinese. Analysis sought to explore how the parallel corpus simplified learner experience in reading, writing and acquiring the Chinese orthographic system, but also to recognize the corpus tool’s potential for guiding students to higher levels of achievement. In other words, the aim was not to simply develop a corpus that would make learning easier, but to develop a tool that would more effectively mediate learner development and enable individuals to function at the top of their ZPD. While the study focused narrowly on the application of a parallel corpus specifically designed for the classroom context in which it was used, the observations and findings have implications for diverse settings in which parallel corpora, or other learning technologies, might be adapted.

The pedagogical application of the parallel corpus in this study was designed to not only provide insight into how the corpus tool aided development, but to provide a functional illustration of both instructor and learner experience in applying the tool. The purpose of this research design was to provide a model of pedagogical application for classroom technology tools. This is not to insinuate that the application observed was perfect, but rather an informed imperfect attempt that provides insight into what aspects of the adaptation of this technology functioned well, and what aspects could be improved. The hope is that future adaptations of parallel corpora (and potentially other technologies) can benefit from the struggles encountered and lessons learned by the instructor and participants in the current study.
Additionally, the novel design of the parallel corpus tool created for this study provides further insight not only into parallel corpus development, but also specifically into the creation of parallel corpora for pedagogical purposes. While there are a few examples of parallel corpora designed for foreign languages classrooms, the majority of extant parallel corpora were created solely for linguistics research, with little thought given to their potential in pedagogy (Bernardini 2003; Lavid, Hita, & Zamorano-Mansilla 2010). Several unique features were implemented in creating the Chinese/English parallel corpus for this project, which – to the researcher’s knowledge – have not been observed elsewhere. One primary innovation is the inclusion of additional language formats intended to aid learner development. Though the parallel corpus is bilingual, the additional language formats allow it to function somewhat similar to a multilingual corpus. Adding the Romanized pinyin allowed for a textual reference learners could refer to when unfamiliar with new Chinese characters. Pinyin is used very frequently in Chinese language instruction, and is regularly incorporated into digital learning platforms. Thus, while the integration of pinyin as a third language format was new in parallel corpus design, it is consistent with other forms of Chinese language learning. Adding the fourth language format, Chinese character + tone mark, functioned to further assist learners by making it possible for them to view only the tone mark to master pronunciation when pinyin was not needed. To the researchers’ knowledge this textual format has not been implemented in other learning materials, and is an innovative feature not just for parallel corpus design, but generally. Beyond these elements, the interactive nature of the parallel corpus, which allows users to click on specific characters to identify tone mark and pinyin, as well as the diverse functions of the query feature also both
provide two further examples of how corpus design can and should consider learner use and experience to enhance potential pedagogical application.

One final purpose identified in the design of the study was the SCT theoretical approach to learning and analysis. The primary instructional approaches observed in parallel corpus pedagogy include reciprocal learning (Hunston 2002; Laviosa 2002), discovery learning (Bernardini 1996), and data driven learning (Frankenberg-Garcia 2005; Wang 2001). None of the extant literature has explored the affects of parallel corpora on development through the SCT lens. SCT provides insight not only into how parallel corpora affect and/or promote user performance, but adds a further perspective of how this technology influences and in some cases transforms learners approach to language learning and conceptual development. This theoretical and analytical framework provides a deeper look at how parallel corpora impact the way individuals learn how to learn, and further enriches the ongoing scholarly conversation of parallel corpus development and pedagogical application.

1.4. Research Questions

To accomplish the stated purposes outlined above, three primary research questions were identified. As outlined, these questions addressed Learner Achievement, Pedagogical Application, and Learner Attitude/Perception:

- Learner Achievement – Will the Chinese/English parallel corpus aid students in learning Chinese?
- Pedagogical Implementation – How do the students and instructor implement the tool in the classroom throughout the semester?
- Learner Attitude/Perception – How do students view the application and implementation of the tool?
The investigation of learner achievement explored several key aspects of how the parallel corpus affected learner experience. This included general consideration of how the corpus aided participants in the acquisition of written Chinese, as well as studying how the four textual formats of the corpus influenced learner experience. Analysis of this question focused both on product and procedure, drawing conclusions of effectiveness based on participant performance with the parallel corpus versus without it, and making observations of how participants’ individual approaches to learning and implementation of the parallel corpus affected their learning outcome and development.

The question of pedagogical implementation is linked closely to learner achievement, as individual application of the tool directly influenced learner outcome. By exploring individual implementation methodologies of both the participants and the instructor, analysis of the this topic led to observations of what approaches were most effective and what uses should be altered, as well as suggestions on how the parallel corpus design could be further improved to enhance user experience. Lastly, addressing learner attitude and perception allowed for an added perspective of participant performance, giving insight into what aspects of the corpus learners found most beneficial, as well as implicating how the implementation of the parallel corpus affected learner motivation in studying Chinese. All of these issues were addressed and analyzed by assessing students’ performance on assignments and exams throughout the semester, and by evaluating screen recordings of participant use of the corpus, learner logs, an end of semester questionnaire, and the instructor’s autoethnography.
1.5. Organization

The organization of the remaining chapters is as follows. The second chapter provides an extensive review of the research done with parallel corpora. The literature review addresses three categories of parallel corpus research. The first category of studies focused specifically on using parallel corpus technology to address specific linguistics topics such as lexical disambiguation and syntax. The second group of research made use of parallel corpora to investigate specific language topics that are challenging for learners, and then recommend methods for pedagogical application – though none of these studies actually applied their research in a classroom setting. The final category includes an extensive review of the published research on pedagogical applications of parallel corpora. Before concluding the chapter, Sociocultural Theory is also discussed in further detail to provide a clear understanding of how the theory understands and analyzes development.

Chapter 3 provides a detailed overview of the design and methodology of the research project. This includes a description of the study and participants, a detailed illustration of how the parallel corpus tool was designed and created, and an explanation of all of the data sources collected and how they were analyzed. Chapter 4 then provides a thorough descriptive analysis of these data sources that is organized both categorically according to type of data source, and longitudinally following the three month-long periods of the study. Chapter 5 reflects on the data analysis to discuss the overall findings of the study and answer the three research questions. This chapter focuses primarily on the discussion of learner achievement and how the parallel corpus tool affected performance and conceptual development, before concluding with the
consideration of both approaches to implementation of the corpus tool and participant opinions and attitudes toward their experience using it. Chapter 6 concludes by discussing the implications of the study results, and suggesting future directions for parallel corpus research.
CHAPTER 2
LITERATURE REVIEW

Although parallel corpora have been used for over two decades in translation studies and comparative language research, their application in foreign language pedagogy is a more recent trend (McEnery & Xiao 2008, p. 21). The relative newness of this field of research means there are limited studies that actually apply parallel corpora in the classroom setting. However, studies do suggest the apparent pedagogical value of parallel corpora, and an increasing amount of work is being pursued in building bilingual and multilingual parallel corpora for general linguistic research and for pedagogical application (Johansson 2009; McEnery & Wilson 2001; Wang 2001; Fan & Xu 2002). Accordingly, the following review of empirical research on the use of parallel corpora in language pedagogy extensively explores the seminal studies that have been published, but has been expanded to include a thorough discussion of the pedagogical implications evidenced in the more diverse body of linguistic research conducted using parallel corpora. The review of parallel corpus research is categorized as: a) studies using parallel corpora to address specific linguistic problems and/or questions; b) research applying parallel corpus technology to address foreign language learning challenges. More specifically, in these studies parallel corpora were used to prepare learning materials and/or address certain learning challenges and then to suggest potential pedagogical applications, but the corpora were not actually used by students; and c) pedagogical applications of parallel corpora in foreign language education. In these studies the language learners used parallel corpora or concordances in learning a foreign language.
The pedagogical application of parallel corpora for the current study is analyzed following the sociocultural theoretical framework in order to evaluate and describe learner development. Therefore, following the discussion of extant research applying parallel corpus research, the sociocultural theoretical model is discussed to provide a clear depiction of the theoretical model applied for the analysis of this study. This section highlights specific aspects of the theoretical framework that relate directly to the use of parallel corpora as a pedagogical medium. Under this subheading, there is also a general discussion of the historical debate on the use of translation for foreign language instruction and more broadly on the use of the L1 in L2 instruction.

2.1. Parallel Corpora in Linguistic Research

As several experts in the field of parallel corpus research are keen to point out, there is an evident lack of research in applying parallel corpora in the classroom (Barlow 2000; Johansson 2009; Wang 2001; Fan & Xu 2002). The foundational pedagogical studies that have been conducted will be extensively explored after first considering the linguistic research that pre-empted the classroom application of this language technology. Parallel corpora have been used in language research for over two decades (McEnery & Xiao 2008, p. 21), and the observations made in the research are directly applicable to foreign language instruction. Therefore, review of this research seeks to not only provide a historical basis but to also highlight several key insights applicable to foreign language pedagogy.

Aijmer and Altenberg (1996) were among the first researchers to advocate and explicitly express the advantages that parallel corpora offer linguistic research. In an
early publication on the topic, the researchers identified four specific uses of parallel corpora in research that are frequently cited by their successors:

1) They give new insights into the languages compared – insights that are not likely to be gained via the study of monolingual corpora.

2) They can be used for a range of comparative purposes and increase our knowledge of language-specific, typological, and cultural differences, as well as of universal features.

3) They illuminate differences between source texts and translations, and between native and non-native texts.

4) They can be used for a number of practical applications, e.g. in lexicography, language teaching and translation. (Aijmer & Altenberg 1996, p. 12).

As explicitly stated, the potential value of applying parallel corpora in language pedagogy (point 4) was recognized early on, but the research that immediately followed focused primarily on the uses of parallel corpora pertaining to linguistic research not pedagogy.

These initial studies focused chiefly on cross-linguistic research. The aligned framework of the parallel corpora provides comparative linguists with an empirical medium for comparing two languages that moves beyond introspection – and allows for the exploration of specific questions of linguistic variance. As Aijmer (2008) highlights, parallel corpora can be used to “disambiguate polysemous items, reveal the degree of mutual correspondence of lexical items in different languages, and uncover cross-linguistic sets of translation equivalents in the languages compared” (p. 99). These functions outline the primary scholarly applications of parallel corpora in the 1990s and early 2000s.

Johansson was one of the pioneering researchers in this field and published several studies outlining his use of parallel corpora in comparing English and Norwegian. In his 1999 study he compares the English verb spend with its Norwegian translation.
This study could have been completed using two monolingual corpora; however, analyzing this question with a parallel corpus provided additional insights to the analysis. In addition to observing the frequency and use of this verb, Johansson was also able to observe how those unique uses were represented in the both languages. Through analyzing the tokens of this verb in both languages Johansson identifies a broader range of functions of the word in English than Norwegian. Whereas the Norwegian tokens of the verb occurred almost exclusively in relation to monetary spending, the English verb included elements “both of passing and using time” (Johansson 1999, p. 52). The primary value and purpose of this study was to identify the different pragmatic applications of the verb, which Johansson recognized as being consistently ambiguous for Norwegian learners of English. The parallel corpus provided contextual examples for the Norwegian learners to observe how the varied meanings of the English verb were typically represented in their L1. Beyond simply addressing this one pragmatic difference between the Norwegian and English, Johansson’s study indicates the potential of using parallel corpus technology to address complex lexical items and disambiguate terms with multiple meanings and uses.

In a second study, Johansson and Hofland (2000) use the same methodology to evaluate the word shall, and the corresponding Norwegian skal. As pointed out at the beginning of the study, Norwegian learners of English exhibit an overwhelming tendency of overusing shall when speaking or writing in English. A statistical analysis of the Norwegian/English parallel corpus represented in Table 2.1 illustrates how the tokens of skal are considerably more frequent than shall, providing insight into why Norwegians would tend to so grossly overuse this lexical item when speaking English.
Table 2.1.

<table>
<thead>
<tr>
<th></th>
<th>E shall</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>skal</td>
<td>36</td>
<td>815</td>
</tr>
<tr>
<td>Original texts</td>
<td>62</td>
<td>561</td>
</tr>
</tbody>
</table>


Johansson and Hofland go beyond this qualitative analysis to provide qualitative data demonstrating the diverse translations of the Norwegian skal in English. Their findings not only elaborate on the intuition of the lexical complexity of skal, but also provide examples of how this word is functionally represented in English – beyond using shall.

Ebeling (1998) provides a third illustration using the same Norwegian-English parallel corpus to explore the meaning and function of the English non-referential there in contrast to the Norwegian determiner det. Ebeling’s analysis reveals not just the statistical variation in function, but qualitative illustrations that demonstrate the pragmatic contrast between the function of there and det in the two languages.

These three studies serve as examples of how parallel corpora enable researchers (and potentially learners) to not just explore the frequency of a concept or word, but to analyze how it is represented in both the source and target language. This allows for the pragmatic disambiguation of concepts and provides a means for discovering a broader range of translation equivalents.

Though the studies considered so far are quite narrow in scope – in that they each only consider a single lexical item - the potential for further application is evident, and substantiated in McEnery and Xiao’s (2008) study of the meaning and function of progressive aspect using a Chinese/English parallel corpus. Their study reveals that the
English progressive has a much wider scope of uses and meanings than does the Chinese (ibid, p. 22). While this observation may seem intuitive to those familiar with both languages, McEnery and Xiao’s use of a parallel corpus confirms this intuition and then identifies the specific differences in how the progressive aspect functions in the two languages. Their research demonstrates how a parallel corpus can be used to disambiguate a challenging concept with a top-down approach and then provide language learners with substance in learning to address this challenging concept.

In addition to helping mediate concepts, parallel corpora can also be applied in broadening conceptual knowledge. As Mauranen (2002) states, “Looking at a parallel corpus, the very first thing that strikes one is the emergence of a much broader range of translation equivalents than in any bilingual dictionary” (p. 168). Parallel corpora provide a much broader reflection of concepts than do traditional sources of language learning and research. Aijmer (2008) continues this same line of thought in stating, “Translations into one or more languages can help us to get a better picture of its meaning and of its correspondences” (Aijmer 2008, p. 95). These two scholars indicate the value of parallel corpora in developing conceptual knowledge. Not only do parallel corpora provide a medium through which concepts can be explored, they also provide a medium through which conceptual knowledge can be expanded.

Of the studies considered so far, the researchers’ primary aims were in using parallel corpora to disambiguate challenging linguistic concepts. In many instances, the topics of research were identified as aspects of language use challenging for language learners to develop (Ebeling 1998; Johansson 1999; Johansson & Hofland 2000). The intent of these studies was not focused on developing pedagogy; however, the potential
pedagogical value is easily identified. In each case, parallel corpora are used to disambiguate the meaning of a term or concept in order to provide a clearer understanding of the issue in both the source and target languages. The results from this research could easily be adapted for the foreign language classroom in order to address the complexity of the respective topics examined as they provide concrete examples and material that are able to represent complete concepts and aid in the navigation of challenging linguistic matters.

2.2. Parallel Corpora in Language Learning

In addition to comparative studies, several scholars have attempted to exploit the niche of parallel corpus research in suggesting new and/or improved methods of foreign language pedagogy. Johansson (2009) lists several applications of parallel corpora relevant to language teaching: developing teaching materials such as textbooks, grammars, and dictionaries; designing syllabuses; training teachers; testing; and in the classroom. According to Johansson, the potential applications of parallel corpora in pedagogy are extensive, and he is not alone in making this claim. The studies considered in the following analysis all conclude with specific recommendations of how pedagogical practices can be improved using parallel corpora.

The transition toward using parallel corpora to improve pedagogical practices has largely evolved over the past decade. Roussel (1991) was ahead of her time, though, in presenting a study that used parallel corpora to analyze English and French tonic auxiliaries. In analyzing this rather complex prosodic linguistic feature, Roussel contends that parallel corpora provide insight into such issues by “showing where problems lie – i.e. where the two languages are furthest apart – and where, by contrast the
two languages work in much the same way” (ibid, p. 81). In her analysis she was able to identify and refine several categorizations of tonic auxiliary function in the two languages (ibid, p. 75). Roussel views these findings as being particularly helpful for teaching advanced language learners. She argues that parallel corpora “can provide an overview of a given language point which may refresh the memory of an advanced learner […] and help things fall into place in his or her mind” (ibid, p. 80). The key point to be taken from her study, and which is cited in nearly every parallel corpus study since, is that parallel corpora allow language learners to see how an item is rendered according to varying contextual elements (Roussel 1991; Wang 2001; Tsai & Choi 2005). In other words, parallel corpora provide a medium through which learners can develop conceptual knowledge of complex linguistic features in diverse contexts, and gradually internalize those concepts.

Aston (1997) also stresses this point in contending that corpora greatly enrich the learning environment by providing students the ability to observe and analyze the “regularities of language use” (p. 63). He takes the claim one step further by asserting that this characteristic of corpora aids students in becoming more independent in their language study by making them less dependent on their instructor. He also argues that corpora reduce teachers’ dependency on textbooks, and allow them “to concentrate on their role as learning rather than language experts” (ibid, p. 63). This line of argument presents the claim that students and teachers are then freed up to become more active and dynamic language learners and explorers. In essence, parallel corpora become an artifact that allows the teacher and learner to cooperatively navigate linguistic concepts in varying contexts and develop together.
Barlow (2000) perhaps summarized this pedagogical function of parallel corpora most concisely as he described them as “on-line contextualized bilingual dictionary[ies]” that provide learners (or teachers) with examples of natural language use (p. 114). In comparing parallel corpora to dictionaries, though, he is quick to point out that dictionaries “lack the richness of context” found in parallel corpora, and also “lack the flexibility afforded by using parallel texts” of being able to search for any linguistic pattern (ibid, p. 114). Barlow’s discussion highlights a couple key pedagogical applications of parallel corpora based upon observations made in prior research that align with the argument made by Roussel (1991). Under the assumption that language consists of “form-meaning links” Barlow (2000) observes that parallel corpora are basically two sets of aligned “form-meaning links”. As the two sets of texts are translations, “the meaning part can be assumed to be approximately the same in both texts. Thus we are able to see how two different language encode equivalent meanings” (ibid, p. 112). In other words, parallel corpora provide more than just ‘meaning in context’ – but provide the researcher/learner with the same meaning in two linguistic contexts.

In Roussel’s (1991) observation of this same idea she argues that parallel corpora could prove helpful for advanced language learners (p. 80). In Barlow’s (2000) expansion of this point, he argues that the dual “form-meaning links” presented by parallel corpora make them ideal for beginning level learners, allowing them to create an awareness of the second language and develop concrete conceptual knowledge. In order for this to be effective, he acknowledges that learners must first overcome the naïve assumption of word-for-word translation equivalence, and instead be taught to look at conceptual meaning. What the arguments of Roussel and Barlow establish, though, is
that parallel corpora have their place in pedagogy at all levels of expertise and learning. This claim has been empirically validated (Frankenberg-Garcia 2000; St. John 2001; Xu & Kawecki 2005), and will be observed in greater detail later on.

In addition to the general pedagogical applications that have been addressed, Bernardini (2003) proposes more specific roles for parallel corpora. Her article describes the process taken in designing and creating the CEXI English/Italian parallel corpus, and she addresses some pedagogical issues that were taken into account during the construction of this corpus. Specifically, she addresses the issue of teaching culturally relevant expressions. She provides an overt lexical example included in the CEXI corpora. In the original English version, Cheltenham - a prestigious British school - is used to emphasize that the subject of the sentence is well off and intelligent. In the corresponding Italian translation, the name of the school is changed to Harrow – another equally prestigious British school that is more familiar to Italians. Bernardini argues that the “observation and discussion of strategies of this kind” is beneficial not just for translators but for all foreign language learners who must learn to navigate not just between two languages, but between two cultures (ibid, p. 534). Accordingly, she claims that the design of the CEXI corpus sought to exploit this inherent feature of translated text by exposing users to cultural comparisons in addition to linguistic juxtapositions.

Hunston (2002) approaches the use of parallel corpora from a different perspective than what has been considered so far. She argues that using parallel corpora is similar to the practice of reciprocal learning. In her words, “Reciprocal learning occurs when two language learners are paired, each helping the other learn their language” (ibid, p. 181). As an example – a native speaker of French would be paired with a native
speaker of English and the two would work to help one another learn their respective languages. Hunston avers that parallel corpora can serve a similar function – either in aiding reciprocal learning, or in lieu of this practice when language partners are not available. In illustration of her claim she provides a personal example in her study of French. She states, “My excitement about parallel concordances is a personal response to seeing English-French concordances, and finding from them enormous amounts of information about French. The use of *dont* before a noun phrase, translated as *including* or *among them*, and the use of *dont* + clause in translation ‘their heads nodding’ and so on were new to me” (ibid, p. 184).

The pedagogical applications of parallel corpora proposed in the studies analyzed thus far spawned directly from the linguistic research of the authors. Other scholars have presented numerous additional methods of applying parallel corpora in pedagogy that have not yet been considered in such detail. Such suggestions include teaching synonymous lexical items, countering prescriptive grammar rules, lexical collocations, and addressing language ambiguity (Pereira 1996; Peters, et.al. 2000; Tsui 1996). Danielsson and Mahlberg (2003) also suggest that parallel corpora could be used in addressing writing problems that arise due to native language influence, as well as being used to “serve as a form of internal differentiation within a group of learners” (para. 17). At this point, it can be concluded that the literature has established several viable applications of parallel corpora in pedagogy, but that there are far more potential applications that have yet to be researched. These studies have made the pedagogical value of parallel corpora evident, but they also serve to emphasize the need for continued research in this burgeoning field.
2.3. Pedagogical Applications of Parallel Corpora

Beyond suggesting potential applications of parallel corpora, a rising body of research has begun to empirically demonstrate the pedagogical value of this technology in language learning. Several of the findings in this body of research concretely affirm the assertions that have already been discussed, while others explore additional avenues of using parallel corpora. The following analysis reflects an extensive review of the work that has been done on applying parallel corpora in the foreign language classroom, including published works as well as conference presentations.

First for consideration is King’s (2003) research, which provides an illustration of Johansson’s (2009) suggestion that parallel corpora be used for material design. In developing his argument for the use of parallel corpora in materials design, King describes the successful account of his colleague Tim Johns in using parallel concordances to develop materials for the teaching of concepts such as “the uses of dont in French, the difference in sense and collocation between schlimm and schlecht in German, and the range of collocates of carry out in English” (King 2003, p. 162). For his own project, King developed a Greek-English parallel concordance to teach the difference between the modals should and would versus will and must. Though there are direct translations of these four modals in both languages, their functions are quite different. King’s study demonstrates that by using parallel concordances students were able to learn “the extent to which natural Greek was produced by translating would as πρέπει (must), and by logical extension, how natural English in certain circumstances uses would to achieve the same pragmatic effect as Greek must” (ibid, p. 162). The use of parallel concordances provided learners with tangible examples and illustrations of
how the modals are conceptualized and function in the two different languages. In King’s words, when using parallel concordances, “we are in effect using translator behavior to hold up a mirror to the differences in realizations of particular functions across languages” (ibid, p. 163). He argues that if instructors focus on the patterns presented by parallel concordances, their application greatly enriches the learning environment of students.

As previously pointed out, Hunston (2002) suggested that parallel corpora could be used to enhance reciprocal learning, or as a second best option when language learning partners are not available. Laviosa (2002) tests this theory with a group of English/Italian reciprocal learners. She used parallel corpora to extract data in both languages in order to “create comprehension and production exercises from which the students infer the rules of usage in each language” (ibid, p. 109). The students were tasked with learning the difference between imperfect, the simple past, and the present perfect tenses in the two languages. They were provided with three sets of parallel concordances for each of the three respective tenses – all in the first person singular. The students were directed to develop a consensual definition of the rules governing the tenses, and then asked to complete a gap filler exercise that was created using additional parallel concordances. Laviosa reported a positive learning outcome, and described the poignant finding that navigation of parallel corpora “reveals precisely the information that the learner needs to acquire in order to establish mental links between L1 and L2 schemas and create new L2 schemas when there is not reciprocity between the two languages” (ibid, p. 110). This point is key in recognizing a unique function of parallel corpora. When used
successfully, they point students toward the precise details needed for optimal development in the language.

Similar to the approach taken by Laviosa, McEnery and Wilson (2001) used the Chemnitz English/German parallel corpus to teach English aspect to German L1 students. The instructors guided the students through a comparison of the two languages and then allowed the students to explore how various combinations and expressions of aspect are communicated in the two languages. At the end of the study, the surveys completed by students affirmed the researchers conclusion that using the parallel corpus proved to be an effective method for developing a conceptual understanding of how aspect functions in English.

Though these first three studies highlight clear advantages of using parallel corpora in the classroom, there are undoubtedly some weaknesses of parallel corpus use that need be considered. As Frankenberg-Garcia (2000) – a leading researcher in parallel corpora - points out, some schools of thought feel languages should be acquired rather than learned (Krashen & Terrel 1983), and since parallel corpora are specialized corpora applicable to only specific groups of learners, their application does not fit in with this line of thinking. Consequently, scholars who adhere to this model have largely overlooked this approach to language pedagogy. Additionally, she states parallel corpora are “impracticable in classes where learners have different native languages” (Frankenberg-Garcia 2000, p. 65). In light of this second observation, it is not surprising that parallel corpus research has not taken off in ESL classroom research, where the classes are frequently composed of students with diverse L1 backgrounds. It should be pointed out, though, that a multilingual parallel corpus could be used in such a setting,
but would undoubtedly present a series of unique challenges. These observations acknowledge certain limitations of parallel corpora application, but do not discount them. As with any pedagogical tool, parallel corpora cannot be applied in every classroom setting, but when their use is warranted they can undoubtedly improve learner development.

Frankenberg-Garcia illustrates the value and effectiveness of parallel corpora in a couple different studies. In an initial attempt, she taught her students to use a Portuguese/English parallel corpus to resolve the misuse of the English preposition *with*. Her students (Portuguese L1 students of English) tended to always translate instances of the corresponding Portuguese preposition *com* as *with*, even though there is not complete correspondence between these two words. In searching the parallel corpus (see Table 2.2), the students quickly discovered that *com* is translated using a variety of English prepositions.

**Table 2.2.**

*Concordance Search Results.*

<table>
<thead>
<tr>
<th>Portuguese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esta foi uma breve panorâmica de uma análise muito completa <em>com</em> cerca de 600 páginas.</td>
<td>This has been a brief overview of some 600 pages of thorough analysis.</td>
</tr>
<tr>
<td>- Parece-me que não estou a compreender o que é que o senhor deputado quer dizer <em>com</em> &quot;mudança de critérios&quot;.</td>
<td>- I am not sure I understand what the honourable Member means by &quot;changing the approach&quot;.</td>
</tr>
<tr>
<td>Por isso, estamos muito animados <em>com</em> o que o Governo búlgaro está a fazer, que é louvável sob muitos aspectos.</td>
<td>So we are very much encouraged by what the Bulgarian Government is doing which is laudable in many respects.</td>
</tr>
</tbody>
</table>

(Frankenberg-Garcia 2000, p. 68).

She found that this exercise not only aided the students in identifying the possible translation equivalents of *com* into English, but also provided them with a clearer understanding of the many diverse functions of the Portuguese preposition (ibid, p. 70). This resulted in them being able to externalize their understanding of this concept through the use of their L1.
In a later study, Frankenberg-Garcia (2005) sought to demonstrate the advantages of parallel corpora in comparison to monolingual corpora in data driven learning exercises. She asserts that parallel corpora can be effectively used by students to answer the questions, “How do you say ______ in L2?” and “Is it okay to say ______ in L2” (ibid, p. 191-2). She states that monolingual corpora can also be used to address the second question, but her study asks the students to determine which resource is most effective for them. Additionally, she states that the first question can be addressed by parallel and not monolingual corpora, and that the corpus approach is more effective than using a dictionary as learners can view how something is said in several different contexts. She argues that this enables learners to not just see how something is said, but provides them with the resources to decide which term is contextually appropriate.

In her study, students were instructed in how to use four different resources to answer the two-abovementioned questions. The resources included a monolingual English dictionary, a bilingual English-Portuguese dictionary, an English-Portuguese parallel corpus, and the BNC simple search monolingual English corpus. After searching out different words and concepts, students then described their experience. One example she provides is of a student who had difficulty understanding the no matter how meaning of the word however. After searching the four different sources, the student gave the following feedback:

The monolingual dictionary (COBUILD, which happens to be corpus based) helped him understand the term, but it didn’t help him see what it meant in his native Portuguese. The bilingual dictionary (Michaelis) explained the meaning of however sought ‘in a more direct way’, but the translation given, por mais que, ‘didn’t quite fit in the sentence.’ The monolingual concordances for however retrieved from the BNC were ‘too many and it took too long to read them all.’ The parallel concordances from Compara were very helpful because they showed him ‘different translations for however in different contexts. There were too many
concordances to read, but a quick glance at the first few was enough’ He concluded that the parallel corpus was the most useful resource, but the bilingual dictionary could save time if it contained some concrete examples (Frankenberg-Garcia 2005, p. 194).

Though this detailed description provides the experience of just one individual, Frankenberg-Garcia notes that the overall preference was for the parallel corpus, and the students’ responses provided specific insight into what aspects of the parallel corpus were most useful. Though it must be acknowledged that monolingual corpora are undoubtedly more appropriate in certain circumstances, Frankenberg-Garcia’s article makes it clear that parallel corpora also have their place in language pedagogy – and that the two types of corpora should be viewed as complementary. She also highlights the potential for using L2-L1 parallel concordances when reading in a foreign language. She claims that such a practice would boost language comprehension and enable learners to navigate more challenging texts that might otherwise be classified as being beyond their level.

The next study highlights Frankenberg-Garcia’s comment about the potential for using parallel corpora in helping learners overcome issues of translation equivalence. In a pilot study, St. John (2001) makes use of a German/English parallel corpus in helping a beginning level German student decipher the meaning of new vocabulary, and formulate grammar rules for himself using the vocabulary. St. John provided instruction in how to use the corpus and then had the student complete a series of nine different tasks. One such task asked the student to look at the German words *sondern* and *aber* using the corpus and try to develop a semantic and/or syntactic distinction between the two (St. John 2001, p. 194). These two German words are often challenging for learners to differentiate as they both translate into English as the conjunction *but*. The student not only produced an accurate description of the distinction between the two, but also picked
up on collocates with both words. The results of this and the other eight tasks led St. John to conclude that the parallel corpus made it possible for the student to more easily and efficiently learn the meaning and proper use of the concepts considered. Ultimately, she posits that parallel corpora not only enable students to understand confusing vocabulary and concepts, but also help them develop an ability to learn how to learn. Her observations strengthen the claims made by Frankenberg-Garcia and demonstrate an effective use of parallel corpora at the very beginning stages of foreign language learning.

Something held in common by the majority of the parallel corpora used in the studies that have been analyzed so far, is that they were initially created for linguistic research and not for pedagogy. In spite of this fact, the above research has evidenced many practical applications of parallel corpora in the classroom. With this tendency in mind, though, one is left to consider the potential of a parallel corpus designed specifically for pedagogy. Lavid, Hita, and Zamorano-Mansilla’s (2010) address this point in their study, in which they design a Spanish/English parallel corpus with the sole intent of applying it in the classroom. Unique to this study is that the intended students were advanced students of linguistics. The authors built their parallel corpus with the goal of creating a resource that their students could use in learning how to apply theory to language analysis. While their article focuses more on the design process versus the actual application – they are keen to point out the successful outcome. They report that students were able to use the corpus to analyze aspectual distinctions between the two languages, practice contrastive analyses, and that they were able to use the corpus to test the students on semantic analysis. While these successes could also be observed using a parallel corpus designed strictly for linguistic analysis, having a corpus designed
specifically for classroom use allowed them to create an interface more accessible to learners, and more easily integrate its application into the classroom setting (ibid, p.145).

Though Lavid, Hita, and Zamorano-Mansilla’s (2010) study focused on a specific group of advanced learners, Wang (2001) is forthright in asserting – as have other scholars – that parallel corpora have their place in helping learners of all levels further develop their knowledge and understanding (Barlow 2000; Frankenberg-Garcia 2005). Additionally, he posits that advanced learners are able to expand their knowledge beyond common meanings of a word and understand how “context in terms of discourse and genre provides clues to the appropriate meaning” (Wang 2001, p. 174). Wang cites personal experience making his claim regarding beginning level learners, but a later study affirms his statement regarding advanced level learners.

Wang (2002) used a Chinese/English parallel corpus to analyze the use of the word xianzai and compared it to the use of now – its English translation. The outcome of this comparison was the observation that the Chinese typically followed a “subject + xianzai” structure, whereas the English had a “subject + be + now” structure. They also analyzed instances where xianzai appeared in the Chinese text, but was omitted from the English translation. The overall analysis provides a more complete depiction of the diverse functions of xianzai and also of the structural differences between the two languages. The presentation of both the structural and contextual elements of the word function in the two languages enables the Chinese L1 advanced level learners of English to develop a more complete conceptual framework in both their target language and their L1.

This general observation is shown to be statistically significant in Fan and Xu’s
(2002) research using a Chinese-English legal parallel corpus. Their pedagogical study was conducted in a university in Hong Kong where students’ L1 is Chinese, but most coursework is completed in English. The legal parallel corpus was created and implemented as a pedagogical instrument to assist the students in developing their English vocabulary and understanding of legal language. The specific goals of their study were to observe how students used the corpus, to get feedback from the students, and to determine whether the corpus could successfully be used to accomplish legal comprehension tasks presented to the students. The students were given two comprehension tasks (in English) that required them to solve legal problems using the parallel corpus. The researchers’ ultimate aim was to discover how useful the corpus was in helping the students comprehend legal texts. After finishing the tasks the students completed a questionnaire that solicited feedback regarding their experience.

Fan and Xu’s analysis focuses more on the questionnaire and how the students used the corpus. They report that 100% of the students reported reading the Chinese text first, and that the majority of the students (about 71%) reported then using mainly the Chinese texts (their L1), but also referring to the English to solve the comprehension tasks. Of particular interest, though, is that over 85% of the students considered using both the Chinese and English (thus the parallel concordances) as being most useful – whereas only 14% reported using only Chinese as being more useful (ibid, p. 55). While the statistical analysis is based on student opinion, and would undoubtedly vary based on individuals and level of language experience, what is noteworthy in this study is that students overwhelming preferred the option of using the parallel corpus. This led the authors to conclude that the use of parallel corpora enhanced the students’ level of
comprehension and learning experience. While some of their methods of analysis could be brought into question, the main point to be taken from their study is indisputable. When parallel corpora are designed and implemented effectively, they not only aid students in development, but students actually prefer using them. This avers that parallel corpora not only aid students in developing conceptual knowledge, but they are potentially more efficient than other methods – as students demonstrate a preference in using them.

Beyond preference, Tsai and Choi (2005) observe that students perform significantly better in lexical development using parallel corpora in comparison to traditional methods. In this study, English L1 students of Chinese were given a list of nine lexical items that were equally divided into three categories. “In terms of the number of English equivalents of a Chinese lexical item, target lexical items were categorized into three different types: 1) one or two equivalents, 2) more than two equivalents, and 3) no exact equivalent” (Tsai & Choi 2005, p. 3). The lexical list was given to both a control group and an experimental group. The control group used traditional methods (dictionaries, textbooks, etc.) to learn the vocabulary, while the experimental group used the Babel Chinese-English parallel corpus, and then both groups were tested on the material.

A statistical analysis of the results of the study indicates that the experimental group participants performed significantly better in categories 2 & 3, and relatively similar in category 1 (Tsai & Choi 2005, p. 4). In a qualitative analysis of the results, it was revealed that in category 2 the control group recalled only the meaning in the textbook, whereas the experimental group remembered more possible meanings.
Additionally, it was observed that the control group tended to “use a direct translation of their L1 to compose a sentence,” while the experimental group did not (Tsai & Choi 2005, p. 6). These findings suggest that those in the experimental groups were developing schemas for the lexical items, whereas the control group was limited to direct translations. Additionally, the lexical items that fell into category 3 required further illustrations and examples as provided by the parallel corpus in order to form new schemas for the target item (ibid, p. 7). The findings of this study poignantly and statistically affirm the arguments made elsewhere (Frankenberg-Garcia 2005; McEnery & Wilson 2001; Fan & Xu 2002), that parallel corpora are able to function as a medium to effectively develop conceptual knowledge of linguistic concepts.

At this point, the pedagogical applications of parallel corpora have been extensively explored, but one final study builds upon the observations made thus far by making use of a multilingual parallel corpus. A multilingual parallel corpus could be employed in pedagogical settings where the students’ target language is the same, but L1’s vary. Alternatively, they could be used in a context where the students all have more than two languages in common. This latter condition describes the environment in which the next study took place.

Xu and Kawecki’s (2005) study was conducted at a university in Hong Kong with a classroom composed of Chinese L1 students with high proficiency in English as their L2. The students were all studying French as a third language, and the design of the multilingual parallel corpus allowed the students to take advantage of both their L1 knowledge as well as their explicit L2 knowledge. The authors found the parallel corpus helpful in aiding students in comprehending linguistic concepts that are often
pragmatically and semantically challenging. As there are greater similarities between English and French than Chinese and French, the authors observe that the trilingual presentation of the texts enabled the students to compare challenging concepts to ideas they had already mastered and learned in English. The unique structure of this corpus and classroom allowed students to make use of their implicit L1 knowledge as well as their explicit L2 knowledge in developing concepts in their L3.

Though the research conducted using parallel corpora is still limited, the studies that have been done clearly demonstrate their pedagogical value. This field of research stands to be enhanced not only by investigating additional pedagogical applications, but also by expanding the theoretical approaches taken by researchers. Of the classroom-based research done using parallel corpora, the most common instructional approaches include reciprocal learning (Hunston 2002; Laviosa 2002), discovery learning (Bernardini 1996), and data driven learning (Frankenberg-Garcia 2005; Wang 2001). Though these approaches have offered valuable insights into the application of parallel corpus, the inherent structure of parallel corpora - which provide conceptual presentations of natural language and take advantage of a learner’s L1 and L2 in developing conceptual knowledge - make them ideal for study within the sociocultural theoretical (SCT) model. The seminal work that has been completed in this field will undoubtedly be enriched by research analyzed from the SCT perspective.

2.3. Sociocultural Theory

The sociocultural theoretical (SCT) perspective posits that learning is a dynamic social activity. The SCT framework is traced back to the work of the Russian scholar Vygotsky in the early 20th century, and has since been further developed in such fields as
psychology, linguistics, and education. Vygotsky depicts cognitive development as a process of internalizing symbolic mediational means through social interaction—transforming “externals into personally meaningful experience” (Lantolf & Thorne 2006, p. 153). This conceptualization emphasizes a functional dialectic between social and individual processes in the co-construction of knowledge (Lantolf 2000a). As outlined by Johnson (2009), SCT asserts that “cognitive development is an interactive process, mediated by culture, context, language, and social interaction” (p. 1). Put frankly, ‘meaning’ and ‘knowledge’ do not simply exist independently, but are socially and culturally constructed and mediated.

This emphasis on mediation in the process of learning and development is one aspect of the theory that has attracted applied linguists. Mediation is viewed as the means through which socially shared activities are transformed into internalized cognitive development and function. Additionally, Johnson and Golombek (2011) point out that it is necessary for learners “to engage in some sort of cognitive struggle” during this developmental process (p. 7). From the perspective of foreign language pedagogy, the role of the educator then becomes a role of assistance through a learning process—presenting learners with cognitive struggles and providing them with the necessary mediation to resolve those struggles. Learning is not viewed as the conveyance of knowledge from the instructor to students, but rather as a cooperative effort in which the instructor actively engages with students and mediates their development.

The mediational role of the instructor is more clearly understood after considering another core concept of SCT – the Zone of Proximal Development (ZPD). Vygotsky states that the ZPD “is the distance between the actual developmental level as determined
by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky 1978, p. 86). Lantolf (2000b) articulates this as being the difference between what an individual can accomplish alone versus what the same individual can accomplish with mediation (p. 52). Thus, it becomes essential for an instructor to approximate an individual’s and/or a group’s ZPD, so that instruction can occur within the ZPD to most optimally assist learners in development.

Also of great importance is the quality of instruction that occurs within the ZPD. One solution within the SCT framework of how to organize and implement instruction is through Concept-Based Instruction (CBI). CBI is an instructional approach that implements the Systemic-theoretical Instruction (STI) model developed by Gal’perin (see Gal’perin 1992 for more on STI). As taken from Gal’perin’s model, CBI stresses that learners are able to achieve greater development using “top-down generalization (transfer)” in which students are taught the theoretical concept and are then instructed and mediated in the navigation of that concept (Arievitch & Stetsenko 2000, p. 79). Serrano-Lopez and Poehler (2008) go on to argue that this characteristic of CBI sets it apart from many traditional pedagogies as it advocates “that learners be introduced to the essential characteristics of their subject matter at the outset of study rather than focusing on narrower sub-topics that might appear easy to learn” (p. 326). In other words CBI (and SCT) posit that learners should be presented with a complete concept (instead of being taught step by step), and then be navigated or mediated through the concept by the instructor via interaction, learning tools, and/or other artifacts. Arievitch and Stetsenko (2000) argue that leaving students to develop their own tactics leads to less productive
and often unsuccessful performance of a task, whereas CBI provides students with the necessary conceptualization and mediation for learning and developing. Basically, CBI provides students with a complete understanding of what a concept is, and then directs them in how to apply that concept in order to learn and develop.

Empirically, it has been demonstrated that the CBI model enables students to develop cognitive theoretical knowledge that can be of value in subsequent learning tasks. For example, Ferreira and Lantolf (2008) demonstrate how CBI improved ESL writing students conceptualization of genre in improving their writing and understanding of composition. Karpov (2005) demonstrates how primary-school children were able to more effectively complete sequence tasks after receiving concept-based instruction versus traditional instruction. In fact ~75% of students in the experimental group were able to apply theoretical knowledge in completing the task, whereas ~20% of students in the control group were able to do so (ibid, p. 191). Neguerela and Lantolf’s (2006) study demonstrates how CBI improved students’ conceptual development and performance in learning Spanish grammatical features such as aspect. Research has clearly indicated the value and efficacy of CBI in learner development and pedagogy.

A key advantage of CBI is that it enables students to not only learn a specific concept, but also to learn to analyze and apply conceptual knowledge in multiple settings. Karpov (2005) claims that CBI impacts “the cognitive development of students such that when they approach new tasks and problems they follow a theoretical and systematic orientation” (p. 189). In addition to applying new conceptual knowledge in diverse settings, is the assumption that learners will apply previous experience and knowledge throughout the learning process. Accordingly, CBI and SCT uphold that foreign
language learning is mediated by an individual’s first language, and their knowledge of how their first language functions (Lantolf & Thorne 2006). Thus, learners are encouraged to apply conceptual knowledge of their L1 in learning a foreign language.

In analyzing the application of the parallel corpus tool in the current study, SCT and the associated concepts of ZPD and CBI are applied in understanding how both students and instructor interact with materials and how development takes place using the parallel corpus as a mediational tool. The inherent nature of parallel corpora enables the user to learn and develop complex linguistic concepts through extant schemas in their native language. While, to the author’s knowledge, the SCT model has not been used to analyze the pedagogical implementation of parallel corpora, it is evident that the theories and ideas posited in parallel corpus research align with the SCT model.

Further, as parallel corpora are composed of source texts plus their translations into another language, rightful theoretical consideration should be given to the use of translated texts in pedagogy. As Frankenberg-Garcia (2004) points out, “It is well documented in the literature that the language of translation is not the same as language which is not constrained by source texts from another language” (p. 225; see also Baker 1996). The “language of translation” (which Salkie (1999) refers to as translationese) should always be understood as a representation of the meaning expressed in the source language and not as a direct equivalent. While this observation is a reality of translation (and therefore an inherent characteristic of parallel corpora), it is undoubtedly a strength that can aid learners in developing their conceptual knowledge of a language.

Translations do not simply expose learners to two linguistic variations of a text, but they provide a written example of how a language expert (translator) chose to
represent the meaning of the source text in his/her translation. As Aijmer (2008) observes, “Translation is one of the very few cases where speakers evaluate meaning relations between expression not as part of some kind of metalinguistic, philosophical or theoretical reflection, but as a normal kind of linguistic activity” (p. 98). Tuğş (2007) also addresses this issue directly in his discussion of parallel corpora, as he states that “the linguistic decisions made by the human translators in order to faithfully convey the meaning of the source text can be traced and used as evidence on linguistic facts which, in a monolingual context, might be unavailable to (or overlooked by) a computer program” (p. 103). In other words, translationese provides a more accurate representation of the meaning of a source text than a computer program or dictionary is able to. Thus, translation texts present learners with a model of a translator’s conceptual knowledge of a language, and parallel corpora enable learners to take advantage of this knowledge in developing their own language skill and conceptual understanding.

In stating that a translation is an expert’s representation of a source text, there is an underlying assumption regarding the quality of a translation and the translator’s level of linguistic expertise. It is therefore necessary to acknowledge “that translation quality is an issue which affects the exploitation of parallel corpora” (McEnery & Wilson 2001, p. 73). Accordingly, a researcher and/or instructor should always consider the translation quality both when constructing and applying parallel corpora. An additional limitation of parallel corpora is corpus size, as there is restricted access to translated texts. For example, there is an evident lack of spoken material, unless interpretation transcripts are included (Johansson 1998). The restricted availability of diverse translated texts, as well as the somewhat extensive processing required to align and to construct a parallel corpus
means that most parallel corpora are relatively small and specialized. Though it is important to be mindful of the limitations inherent to parallel corpora, both McEnery and Wilson (2001) and Johansson (1998) stress that these weaknesses do not invalidate their research and use, but rather emphasize the importance of careful corpus and research design in exploiting parallel corpora.
CHAPTER 3
RESEARCH DESIGN AND METHOD

For this study a new Chinese/English parallel corpus was designed and constructed specifically for the purpose of language learning. The primary aim in creating this parallel corpus tool was to develop a resource that would make written texts more accessible to language learners, and improve upon current approaches to written language instruction. Two primary challenges faced by language learners in written language acquisition include reading comprehension and sentence composition and construction. These tasks are particularly difficult to address at beginning levels of language learning because learners may not have developed adequate vocabulary or functional knowledge of the language. The parallel corpus tool designed for this project addresses these and other tasks of language development by making written language more accessible to language learners at all levels.

While reading comprehension and writing development in any foreign language can be challenging, the unique and complex structure of written Chinese presents learners with additional obstacles. The Chinese orthographic system is composed of thousands of individual characters, with estimates suggesting that an individual must know 3,000 – 4,000 characters in order to read general texts such as newspapers (Norman 1988). The large amount of characters that one needs to acquire makes learning written Chinese especially challenging for beginners since the Chinese writing system is not based on an alphabet, and students cannot simply sound out words or phrases. In alphabetically based languages, learners can sound out words and read through entire documents after mastering the alphabet and phonology associated with the language. This provides
learners with the benefit of aural recognition as well as context to derive the meaning of
unknown terms or phrases and to decipher textual meaning. Though being able to read
through a document and understand it are two different things, it is important to
remember that learners of Chinese face additional obstacles before even being able to
read through a text. Consequently, Chinese language learners have less access to
contextual information and/or the ability to draw on aural recognition when encountering
new written texts than do learners of alphabetic languages.

Another unique feature of Chinese is the composition of words. A Chinese word
can be composed of one, two, or even more characters, and many characters can have
several meanings. For example the character 会 (huì) generally means to be able to, but
has additional uses. It can also be combined with other characters to form two-character
words such as 社会 (shèhuì), which means society. Though the same character 会 (huì)
appears in both words, the two words have very different meanings and function
differently within sentences. So an individual who had learned the word 会 (huì) but not
社会 (shèhuì) would be easily confused and likely completely misunderstand a text that
contained the latter term. This homonymic feature of many Chinese characters presents
learners with several challenges. At one level, simply confusing word order within a
sentence could result in an alternate meaning being expressed in the text. At another
level, remaining unaware of multiple meanings and functions of terms greatly limits a
learner’s ability to effectively function and communicate within a foreign language.

These inherent challenges to learning Chinese orthography are specifically ad-
dressed and made more accessible through the use of parallel corpus technology. These
explicit benefits of parallel corpus technology add to the already established claim that
using parallel corpora in language teaching enable students to conceptualize the target language through schemas in their L1. The corpus tool designed for this study improves upon this inherent characteristic of parallel corpora by including added features that enable learners to read through Chinese texts fluidly, and not be limited by the characters they may not know yet. Not only does this allow individuals to benefit from aural recognition and context clues, but it also makes more advanced texts accessible to learners. Additionally, parallel corpora also provide learners with an efficient method for addressing the challenge of polysemy, homonymy, and multiple functions of characters. Corpora, in general, work well in addressing polysemy, but parallel corpora assist learners in more readily comprehending concepts through their first language.

Specifically, this study integrates the use of the parallel corpus tool to address three primary research questions: 1) Learner Achievement – Will the Chinese/English parallel corpus aid students in learning Chinese? 2) Pedagogical Implementation – How do students and the instructor implement the tool in the classroom and their study throughout the semester? And 3) Learner Attitude/Perception – How do students view the application and implementation of the tool? A detailed discussion of these research questions, and how they were assessed and evaluated is found in section 3.3. Data Analysis.

3.1. Context of the Study

3.1.1. Setting

The study was conducted in two beginning level high school Chinese language classes at a public high school in central Pennsylvania. The two classes were taught by the same instructor (researcher) and data was collected over a three-month period during
the spring 2013 semester. The two classes had similar levels of language instruction prior to and during the study, but were scheduled differently. One class (Group A) met three times a week on Mondays, Wednesdays, and Fridays, whereas the second class (Group B) met daily Monday through Friday. Both groups had open access to technology in their classroom setting. Group A had open classroom access to computers and the Internet that the course provided as a learning resource. Group B also had access to the Internet and computers, but additionally each student in this group had a school-issued iPad2. The computers and iPad2s were set up with screen-recording software that was used to record students’ use of the parallel corpus and other online resources during their classroom time. Students also had open access to the parallel corpus outside of class, and use of the tool was reported in learner logs kept throughout the duration of the study.

3.1.2. Participants

Twelve high school students enrolled in beginning level Chinese participated in this study. Consent was obtained from both students and their parents in order to take part in this research project. Additionally, appropriate permissions were obtained from the school district in order to conduct this study. Participants ranged in age from fifteen to eighteen years old and in grade from ninth to twelfth grade. All students were enrolled in their second semester of Chinese, and had been with the same instructors throughout the duration of their study. The Chinese courses had been team-taught by two instructors: one instructor being a native Chinese speaker and the other being a native English speaker with advanced Chinese proficiency. During the study the courses were taught exclusively by the latter instructor. The participants came from two different
classes held on separate campuses (nine from group A and three from group B) and did not include heritage learners. All students enrolled in the classes chose to participate in the study. One student’s first language was Korean, but with advanced proficiency in English as an L2. All other participants shared English as their first language. Pseudonyms are used to identify students in the study in order to maintain anonymity.

3.1.3 Course details

This study took place over a three-month period in which all participants in both classes received the same treatment. The course offering was an elective and students received elective credit for the course, but this did not count as foreign language credit as the school district has not yet moved to approve a credit bearing Chinese language curriculum. The study began approximately two months into the semester and during the first month of research the course continued unaltered, with the exception that participant data (assignments, exams, learner logs, etc.) were collected for analysis. Assignments were completed every other class period (once or twice per week), and there was a unit exam approximately every two weeks. The parallel corpus tool was introduced at the beginning of the second month of the study. At this time, participants were taught how to use the tool and were encouraged, but not required, to use it in their study and class work as a resource for looking up new lexical items as well as examples of sentence construction. This time period was used as an instructional practice period in which the students became comfortable with using the tool and shared, through learner logs, their experiences in adapting it and the extent to which they used the tool. By the third and final month of the study, the tool was completely integrated into the classroom as a regular part of participant activity during all exams and for all in-class assignments.
Participants had continual access to the parallel corpus tool during class time, and were encouraged to use it as needed even if the focus of the class did not specifically require its use. Additionally, students could access the parallel corpus tool outside of the classroom as it was provided open-access on the Internet.

The course curriculum was set up following a textbook that the high school had purchased for students, Learn Chinese with Me (跟我学汉语) (Chen 2009). In prior years the high school had offered Chinese to interested students and provided them with a peer instructor who was a Chinese-English bilingual. After arranging for an instructor to teach the course the decision was made to continue the curriculum based on the textbook. Accordingly, the themes and lessons designed and integrated in the classroom during this study originated from the topics in the textbook. Classroom instruction and activities were primarily content-focused based on textbook topics, and grammar lessons were integrated when helpful for learning functional expressions. Typically, two class periods were taken to complete one lesson topic from the textbook. The first class was instructional, and the second class was practical. During the first class period on a textbook topic new vocabulary, grammar, and associated topics were introduced and discussed and an assignment given. The assignment would often be started in class and then taken home if necessary for completion. During the second class period on a topic the assignment content would be discussed and the new material would be practiced through varied activities. After completing two textbook lessons (after four class periods), there would be a review activity and assignment followed by an exam during the next class period. In addition to the language instruction, the Friday class period was set aside as a separate cultural experience day. Participants joined students from different
Chinese levels to discuss and experience various topics of Chinese culture. These class periods were separate from the regular curriculum and were not included in this study.

Table 3.1 depicts group A’s course schedule. Group B’s schedule was identical, but with different days and dates. As can be observed in the table, approximately two lessons were covered each month, meaning one exam was given during each of the three phases of the study.

**Table 3.1.**

*Course Schedule.*

<table>
<thead>
<tr>
<th>Class Date</th>
<th>Plan</th>
<th>Homework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday, 11 March</td>
<td>Lesson 13 – Review Vocabulary, Individual Student Conferences, In-class Workbook Assignment</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 13 March</td>
<td>Lesson 13 – Finish Lesson 14 Introduction</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 18 March</td>
<td><em>Snow Day</em></td>
<td></td>
</tr>
<tr>
<td>Wednesday, 20 March</td>
<td>Lesson 14</td>
<td></td>
</tr>
<tr>
<td>Monday, 25 March</td>
<td><em>Snow Day</em></td>
<td></td>
</tr>
<tr>
<td>Wednesday, 3 April</td>
<td>Review Activity</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 8 April</td>
<td>Exam</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 10 April</td>
<td>Outside Activity (Internet Down)</td>
<td></td>
</tr>
<tr>
<td>Monday, 15 April</td>
<td>Story/Project, Lesson 15</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 17 April</td>
<td>Lesson 16</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 22 April</td>
<td>Review Activity</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Wednesday, 24 April</td>
<td>Exam</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 29 April</td>
<td>Lesson 17</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 1 May</td>
<td>Lesson 17</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 6 May</td>
<td>Lesson 18</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 8 May</td>
<td>Lesson 18</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Monday, 13 May</td>
<td>Story/Project</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 15 May</td>
<td>Review Activity</td>
<td></td>
</tr>
<tr>
<td>Monday, 20 May</td>
<td>Exam</td>
<td>Assignment Due</td>
</tr>
<tr>
<td>Wednesday, 22 May</td>
<td>Year Review</td>
<td></td>
</tr>
<tr>
<td>Wednesday, 29 May</td>
<td>Final Project/Exam</td>
<td></td>
</tr>
<tr>
<td>Monday, 3 June</td>
<td>Course Reflection, End of Semester Questionnaire</td>
<td></td>
</tr>
</tbody>
</table>
Lesson topics included age, animals, locations, directions, family, and occupations. Activities were designed to promote lexical acquisition of written Chinese characters and spoken interaction on the topics. Though the topics were determined by the textbook-based curriculum, the classroom instruction was not limited to the textbook but simply used it as a topic guide. In other words the lessons and content covered in class was based on the topics, but designed by the instructor/researcher.

3.2. The Parallel Corpus Tool

3.2.1. Interface and Design

The Parallel Corpus Teaching Tool (Bluemel 2013) used in this study was designed specifically for pedagogical application in the FL classroom setting. It was created as an interactive web-based tool formatted for both computer and tablet access. As mentioned, one of the study groups used school-issued iPad2s, and both groups had open access to computers, iPads, or both during class time. The participants also had open access to the tool and could access it from home using personal devices.

The creation of a new corpus tool, instead of adapting an extant one, was pursued for several reasons. First, several researchers have established the need for more and better-developed parallel corpora (Fan & Xu 2002; Johansson, 2009; Laviosa 2002; Wang 2001). Second, developing a new corpus tool allowed for innovative functions that more precisely address the challenges of Chinese orthography. Third, research has demonstrated that designing a parallel corpus specifically for pedagogy improves students learning experience by allowing for an interface that is more accessible and easier to integrate into the curriculum (Lavid, Hita, & Zamorano-Mansilla 2010). The majority of parallel corpora currently available have been created for linguistic research,
without considering the possibility of pedagogical application. The design and features of the Chinese/English parallel corpus created for this study were greatly influenced by the intent for the pedagogical application of the tool.

Xu and Kawecki’s (2005) study using a trilingual English/Chinese/French parallel corpus suggests the value of presenting parallel corpora in more than just the standard bilingual format. Though the parallel corpus tool created for this study is bilingual (Chinese/English), it includes texts in four language formats: Chinese characters, Chinese characters + tone marks, pinyin, and English. Just as the students in Xu and Kawecki’s (2005) research were able to use both English and Chinese in learning concepts in French, the design of the Parallel Corpus Teaching Tool enables students to use tone marks, pinyin, and English as aids in learning Chinese characters. In order to realize the significance of this design structure, it is necessary to first consider some basic elements of Chinese.

Though Chinese is written using characters, a corresponding system known as pinyin has become the standard writing system for transliterating Chinese characters by using the Roman alphabet. It is used both by FL learners of Chinese as well as native speakers of the language. The pinyin system allows for the alphabetic representation of characters, which can aid in reading, understanding, and typing Chinese. Another pertinent feature of Chinese orthography is that, generally, each character corresponds to one syllable, and every syllable/character has a tone mark. Chinese is a tonal language, and the tone associated with each syllable/character functions to indicate the meaning. Mandarin Chinese has four tones, plus a fifth neutral tone, making it possible for a syllabic utterance such as ma, to have five different possible meanings based upon the
tone. It is therefore imperative that utterances in Chinese are spoken with the correct tone, and that students learn the correct tones associated with characters and meanings of words.

To aid in the mastery of tones, tone marks are used as part of the pinyin system to clearly demarcate tone. Tone marks appear in two forms, either as numbers following a syllable or, more typically, as diacritics written above the syllable, as illustrated in Table 3.2. In studying Chinese, learners typically begin by first learning the pinyin system and tone marks before moving onto characters.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Tone indicated by Number</th>
<th>Tone indicated by Diacritic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Tone</td>
<td>ma1</td>
<td>mā</td>
</tr>
<tr>
<td>2nd Tone</td>
<td>ma2</td>
<td>má</td>
</tr>
<tr>
<td>3rd Tone</td>
<td>ma3</td>
<td>mā</td>
</tr>
<tr>
<td>4th Tone</td>
<td>ma4</td>
<td>mà</td>
</tr>
<tr>
<td>Neutral Tone</td>
<td>ma</td>
<td>ma</td>
</tr>
</tbody>
</table>

As the parallel corpus tool was designed, the function of pinyin and tone marks in learning written Chinese was integrated into the tool by aligning all the corpus texts in these four language formats (Chinese characters, Chinese characters + tone marks, pinyin, and English). Figure 3.1 depicts these four different formats. The Chinese character and English formats lie at opposite ends of the spectrum, and require little discussion as they simply represent the two languages. The character + tone marks and pinyin formats function as a form of literacy learner language to assist learners in the acquisition of characters. While the pinyin representation of any character is available in most Chinese language technologies, the Character + tone mark format is a novel ap-
approach, which, to the researcher’s knowledge has not been previously implemented elsewhere. These two additional formats specifically address the challenges learners face in being able to read complete texts, and are illustrated below.

**Figure 3.1.**

*The Four Text Formats.*

Though the texts are aligned in all four of the language formats, students are able to select which formats are visible. The goal is for students to use the least amount of mediation necessary to read and understand texts. Ideally, they would be able to read the Chinese characters, but if additional information is needed, they can then add tone marks to help with pronunciation. Thus, as will be demonstrated in the data analysis, the tool aids learners in written language acquisition by enabling them to comprehend texts by using additional formats as necessary, but also places an emphasis on character acquisition by presenting learners first with just the characters. This contrasts with a majority of beginning level learning materials that always juxtapose characters with pinyin, enabling students to completely ignore characters. In using the parallel corpus, learners can click on a character to add pinyin, but the primary focus is placed on character acquisition, only making pinyin visible as needed. Last, the English translation is also available for words or characters they are not able to derive the meaning of using the other three textual formats.
The aim of this aspect of the design was to provide a learning environment in which students focus on character acquisition, and tone marks and pinyin are used as a form of literacy interlanguage to assist them in the ultimate goal of learning characters. Systematically, the corpus did not have a way to enforce usage in this manner, but classroom instruction on proper usage as well as screen-recordings of students’ use of the tool ensured its application for the intended purpose. The corpus allowed users to view entire texts, or they could also query a specific term or phrase using Chinese characters, pinyin, and/or English. A character input system was necessary to search for Chinese characters, whereas the tool included a feature that not only enabled, but also required, students to input the correct tone mark in order to search for terms using pinyin. When a student types a pinyin letter that could possibly contain a diacritic, the possible options appear on the screen and require the student to select the correct one.

The ultimate success of the application of the Chinese-English parallel corpus in the classroom is dependent upon the functionality and accessibility of the tool itself. Most research on corpus design focuses on the balanced composition of material used in compiling the corpus (Bernardini 2003). While this remains important in the construction of any corpus, the intended pedagogical application of the corpus constructed for this project implies the need to pay special attention to interface design. Hadjerrouit (2010) established a conceptual framework for using web-based learning resources (WBLRs), and empirically established a list of key criterion to consider. Hadjerrouit lists three primary categories to consider: WBLR features, usability criteria, and context of use/evaluation; then lists 26 key points to consider under these three headings (ibid, p. 61). Hadjerrouit’s (2010) WBLR Framework is provided in the
Appendix and depicts the detailed conceptual framework that was used as a guideline in creating and designing the user interface of the parallel corpus tool. While Hadjerrouit’s conceptual framework was used as a standard, the final details of the design were ultimately worked out in consultation with the third party software design expert who was employed to create the tool.

3.2.2. Corpus Materials

The approach taken in compiling texts into the digital parallel corpus tool is somewhat unconventional, as the intention was to maximize the pedagogical value of the tool. The corpus is composed of bidirectional texts (Chinese L1 to English L2 translations and English L1 to Chinese L2 translations) that are functionally accessible to students, meaning the students are familiar with the content addressed in the included texts. The texts included in the corpora were selected according to two criteria: 1) the content of the texts related to the content covered in the course curriculum, and 2) the texts were challenging in that students did not necessarily know all words or characters, but appropriate for students’ language level. The basic set of texts included the students’ textbook, *Learn Chinese with Me* (跟我学汉语) (Chen 2009) along with several supplementary articles and books, which were selected by the instructor. The supplementary texts used in the corpus included bilingual short stories (Hou 2006; White 2004), bilingual textbook articles (The Overseas Chinese Affairs Office 2007a; 2007b; 2007c), bilingual published speeches (Xu 2011), and bilingual online articles (National Foreign Language Center 2010).

While only a small subset of texts was initially made accessible through the corpus tool, the course instructor was able to add, change, or alter source texts throughout
the semester. This means that the corpus size was continually expanding, and by the end of the research period the corpus contained 45 texts aligned in Chinese and English, of which approximately 18 were from the textbook and the remaining were from the selected supplementary articles and books. These 45 texts contained 26,563 words (6,512 Chinese words or tokens (11,039 Chinese characters), 6,512 Chinese characters + tone mark tokens, 6,512 pinyin tokens, and 7,057 English tokens). By enabling the instructor to control the data that students were able to access, the goal was not to limit their exposure, but rather to insure that students were able to access material that is consistent with their learning level and with the content being covered in class, and thus be guided to optimal development within their ZPD. By doing so, the instructor is able to add more challenging texts as more vocabulary and content is learned throughout the semester, and hopefully provide texts that continue to challenge students in their development. This included adding texts that contained new lexical items as well as gradually adding texts of greater length.

All texts used in the corpus were aligned at both the word and the sentence level. These two levels of alignment allow for different functions within the tool. First, the word alignment allows students to search for individual words, and have the corresponding translation equivalents highlighted in the corresponding text formats (see Figure 3.2). The figure depicts the search term 马 with its corresponding pinyin mā and English horse all highlighted in green in the text. Additionally, an interactive feature of the tool allows the student to view the text in the character with tone mark format after clicking on the Chinese character.
The texts were also aligned at the sentential level for two reasons. First, the fluidity of text sources required a more standardized system of alignment. In other words, sentence level alignment provided an organized schematic in which to store and organize source texts. Second, sentence alignment forces students to analyze the structure of the entire sentence, and observe how the two languages differ grammatically. This is accomplished as the query results for any given term or phrase are displayed within the sentence they appear in, meaning that all results are displayed in sentence format. By having both word-level and sentence-level alignment students are also able to identify specific terms and then compare how those terms function within the two different languages.

Chinese/ English bilinguals reviewed all source texts included in the corpus in order to insure the accuracy and validity of the translation material. The researcher is an
English L1 advanced Chinese L2 speaker and reviewed all data sources used for the corpus. Additionally, a Chinese L1 advanced English L2 speaker was also employed to review all texts before including them in the corpus.

3.3. Data Analysis

3.3.1 Data

Several varied data sources were collected to analyze the pedagogical implementation of the parallel corpus tool and the instructor and learners’ experiences. These sources include: 1) Screen Recordings, 2) Learner Logs and a Questionnaire, 3) Instructor Autoethnography, and 4) Course Material. These varied data sources were collected throughout the duration of the study and then compiled and analyzed both individually and collectively. A brief overview of data is given in Table 3.3, and each of the data sources is described in further detail below.

Table 3.3. Data Type and Description.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screen recordings</td>
<td>Video recordings of participant computer and/or iPad screens while completing in class assignments and exams</td>
<td>48 video recordings</td>
</tr>
<tr>
<td>Learner Logs</td>
<td>Guided reflection prompts of learner experiences in challenges on individual assignments and exams</td>
<td>111 Submitted logs</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>End of semester questionnaire completed by participants</td>
<td>12 questionnaires</td>
</tr>
<tr>
<td>Instructor Autoethnography</td>
<td>Journal entries of instructor’s experience in preparing for and reflecting on implementation of tool</td>
<td>22 Entries</td>
</tr>
<tr>
<td>Course Material</td>
<td>Assignments, exams, and other material (see Table 3.4 for complete list of details)</td>
<td>147 Documents</td>
</tr>
</tbody>
</table>
3.3.1.1. Screen Recordings

Screen recordings were taken of participants’ use of the tool during in-class review assignments and during all exams. These recordings lasted the entire duration of the class period and captured all details of the students’ use of the computers and/or iPads. The recordings included only the visual screen captures of the participants’ screens, no audio data was obtained. All recordings were stored and not reviewed until after the completion of the study and the course, thus screen-recordings were not used to evaluate student performance for the course and in no way affected student grades. The purpose of the screen recordings was to document exactly how each participant chose to implement the corpus tool in completing their work. Additionally, the recordings depict what features of the tool the participants used most frequently, and which features the participants seemed to ignore.

Participants in Group A only had access to computers in the classroom. Thus, all in-class activities and screen recordings were completed on either MacBook laptops or Mac desktop computers. Screen recordings were made using QuickTime Player Version 10.3 software. This software comes pre-installed on Apple computers, and was easily accessible to complete all screen-recordings. At the beginning of the class period students would begin their screen recordings, and at the end of the class period the instructor/researcher would save and store all recordings.

The students in Group B each had school-issued iPad2s that were used to complete in-class activities and assignments. Screen recordings of the students’ iPad2s were made using both Reflector Version 1.5.4 and QuickTime Player Version 10.3 software. This Reflector software was installed on the instructor/researcher’s MacBook
Pro computer. This software reflects the students’ iPad screens onto the researcher’s MacBook Pro, thus displaying all iPad screens on the laptop screen. QuickTime Player was then used to record the researcher’s computer screen.

3.3.1.2. Learner Logs and Questionnaire

Throughout the entire school year the students were tasked with completing learner logs as part of the course curriculum. These learner logs were used to help bring an explicit awareness to the students of their individual study habits and what worked most effectively for them. The students were provided with several questions prompting responses regarding what resources they used and how they used them, what challenges they faced, and what assistance they felt they needed in the future. Additionally, students were also encouraged to journal any additional thoughts they had about their learning experience.

Prior to the study, this task was less formal and students were encouraged to always respond to one or two of the prompts for each assignment and exam. During the study, students were asked to respond to all of these questions. A handout with these prompts was attached to each assignment and exam that was given to the students. A copy of this handout can be viewed in the Appendix, Learning Log Prompt. The handout includes six different prompts that the students responded to.

As mentioned, the learner logs were required of the students throughout the entire academic year, but were more formalized during the duration of the study. The learner logs served two primary functions: 1) they provided detailed feedback from the participants during the study regarding their language learning experience both with and without the parallel corpus tool. 2) Participants became aware of how they learn a
language, and what language learning strategies and resources worked best for them. At the end of the academic year, students were provided with a summarized analysis of their individual learner logs that highlighted what strategies they felt worked best for them, which were less effective, as well as suggestions of new approaches to take as they continued their study of the language. While this information was obviously very useful in analyzing participants’ adaptation of the parallel corpus tool, the initial goal of this task and data source was to assist learners in their language development by teaching them how to learn a language.

Additionally, students completed a questionnaire at the end of the study. This questionnaire was similar to the learning logs, but more extensive. Whereas the learner logs prompted participants to evaluate their learning experience on one specific task or assignment, the questionnaire prompted them to analyze their overall learning experience through the duration of the study. Similar to the learning logs, the questionnaire prompted narrative responses from the participants. However, it also included a section that asked participants to use a Likert scale to rate the parallel corpus tool. The complete form is included in the Appendix, End of Semester Questionnaire.

3.3.1.3. Instructor Autoethnography

Similar to the learner logs completed by the participants throughout the study, the instructor/researcher also completed a guided journal. For each class period, the instructor completed an autoethnographic journal entry that analyzed his preparation and experience in teaching the class and in using the parallel corpus tool. The journal entries were guided responses using a prepared handout with ten questions. The handout can be viewed in the Appendix, Autoethnography. As stated directly on the autoethnography
handout, the purpose of the instructor journal was to “include entries on how the instructor prepares for each class and prepares to use the tool in class, as well as post-instructional accounts that detail how the class (and daily implementation of the tool) went”. The journal entries were completed during all three months of the study, even though the parallel corpus tool was only implemented beginning the second month. Thus, the entries from the first month of the study simply disregarded the prompts relating specifically to the use of the parallel corpus tool.

The intent of this data source was to provide insight into the instructor’s experience in implementing this new technology, by highlighting both successes and failures. The hope was that this individual case study would provide a model for more successful future pedagogical adaptations of parallel corpus technology, or other language learning technology.

3.3.1.4. Course Material

All course assignments, exams, and materials were collected during the duration of the three-month study. Participants’ performance on the assignments and exams were used to assess and analyze student understanding and development. The assignments and exams included lexical acquisition tasks, reading comprehension tasks, as well as writing tasks.

As an example of one of the students’ learning activities, after learning Chinese kinship terms, the participants were tasked with using the parallel corpus tool to read a letter in which the author, Xiūming, described the members of his family. The letter included many of the new vocabulary terms recently covered in class and additional unfamiliar terms. In addition to reading and explaining the meaning of the letter through
a written prompt, participants were asked to identify new lexical items and then search those items in the parallel corpus in order to find additional examples of their use and then describe the meaning of the terms and how they were used. Last, participants completed a separate writing task in which they had to use the new lexical items they had individually identified in order to write a response letter to Xiūmíng describing their own family. Data collected from this particular task included the written assignment itself, the associated learner log, as well as screen recordings of students’ use of the corpus tool while working on the task in class.

As previously stated, assignments such as this one were given every other class period, and one unit would be completed approximately every two weeks. At the end of each unit students were given a review assignment, which was then followed by an exam. Thus, the course material analyzed for this study includes several class assignments, unit review assignments, and exams. Table 3.4 provides further details regarding the sources of course material that was completed and analyzed for this study.

Table 3.4.

<table>
<thead>
<tr>
<th>Course Material</th>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assignments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>You’re Going to China!</td>
<td>Students planned an imaginary 10 day trip to China</td>
<td>11 Submitted Assignments</td>
</tr>
<tr>
<td>*Lesson 13 - Age and Activities</td>
<td>Vocabulary exercises, writing sentences with new vocabulary items</td>
<td>11 Submitted Assignments</td>
</tr>
<tr>
<td>*Lesson 14 - Animals</td>
<td>Vocabulary exercises, in-class interview about classmate’s animals, writing assignment describing a pet or animal</td>
<td>9 Submitted Assignments</td>
</tr>
</tbody>
</table>
*Zodiac Story

Students read the story of the *12 Zodiac Animals* then completed a question and short answer worksheet and summarized the story

10 Submitted Assignments

*Lesson 15 - Hometown and Countries

Vocabulary exercises, writing sentences with new vocabulary items

10 Submitted Assignments

*Lesson 16 – Addresses, Directions, and Food

Vocabulary exercises, completed dialogs with new vocabulary and practiced in class

11 Submitted Assignments

Webquest

Web search assignment that guided students in learning about Mao Zedong and the Cultural Revolution

12 Submitted Assignments

*Lesson 17 – Family Members

Students read a letter from “Xiūmíng” that described his family, then wrote back to him describing their own family

10 Submitted Assignments

*Lesson 18 - Occupations

Students wrote an essay describing their parents, or other family members, occupations, and described their dream job

8 Submitted Assignments

*Review

Vocabulary exercises, reading comprehension exercises, sentence composition

10 Submitted Assignments

Presentations

What Interests You About China

Students gave a 3-4 minute presentation in Chinese about what intrigues them most about China. Notes and prepared scripts for the presentations were collected

9 Completed Projects

Exams

*Lesson 13 & 14 Exam – Age and Animals

Vocabulary, questions and short answer sentence composition, describe images using complete sentences

12 Exams
3.3.2. Analytical Framework/approach

As stated, the primary aims of the current study were to analyze the pedagogical value of the Chinese/English parallel corpus tool, and to detail how both the students and the instructor apply, implement, and view the tool. The three primary issues considered in this study include: 1) Learner Achievement – Will the Chinese/English parallel corpus tool aid students in learning Chinese? This question encompasses both learner achievement in the acquisition of written Chinese as well as considering whether the unique four textual formats of the corpus assist learners in more efficiently learning the language. 2) Pedagogical Implementation – How do the students and instructor implement the tool in the classroom throughout the semester? And 3) Learner Attitude/Perception – How do students view the application and implementation of the tool? The varied data sources described above were all analyzed both individually and then collectively to address the three primary questions. The Sociocultural theoretical (SCT) model was applied to analyze the data for evidence of development.

SCT recognizes learning and development as a dynamic social activity and asserts, “cognitive development is an interactive process, mediated by culture, context,
language, and social interaction” (Johnson 2009, p. 1). In other words, a participant’s (or individual’s) knowledge and understanding of a topic are not independent categories, and cannot be categorized as such, because knowledge and meaning are both socially and culturally constructed and mediated. In the current study, the primary analytical emphasis is on understanding and describing how the parallel corpus functions as a mediational tool in aiding learner development and knowledge of written Chinese. Thus, the data is not simply analyzed to find evidence of student improvement, but rather to identify if and how the parallel corpus tool was able to mediate learner development. The data was therefore analyzed for evidence of participants’ learning and internalizing lexical, grammatical, and other functional linguistic constructs. How this was approached to address each of the three research questions is discussed in greater detail below.

*Learner Achievement* - All of the structured tasks completed throughout the semester were graded, and students’ performance on tasks was taken as a preliminary measure of achievement. Though quantitative analysis of the results was evaluated and used as a preliminary measure, it is acknowledged that the small group size limits the researchers ability to generalize the statistical findings of the study. The primary focus of analysis, though, was on a qualitative analysis that explores how students apply their knowledge of new language features in completing the tasks. In Tsai & Choi’s (2005) study they observed that students using a parallel corpus were able to recognize multiple functions of specific lexical items, as well as develop more complex sentences in their writing than students who did not use a parallel corpus. The qualitative analysis of student work considered similar findings in reviewing the lexical acquisition tasks,
reading comprehension tasks, and writing assignments. The analysis of the assignments was completed longitudinally and follows the progress of each individual student throughout the semester – looking at assignments and exams for performance and cross-referencing performance with data from the screen recordings that evidenced students understanding (or lack thereof) of individual concepts and terms.

**Pedagogical Implementation** was analyzed using the screen recordings, learner logs, and the instructor’s autoethnography. The screen recordings of how students use the parallel corpus in completing the unit projects were evaluated individually to determine exactly how each student used the corpus tool. Specifically, the researcher looked to see which textual formats the students relied on, and sought to establish any patterns of how the students adapt the tool to meet their needs in completing the task. The findings of the screen recordings were then evaluated in context of student performance on assignments and exams in order to establish correlations between student learning outcomes and how they use the tool. In other words, the data was analyzed to see if students who performed well on assignments and exams exhibited unique patterns in how they used the parallel corpus tool.

Additionally, the learner logs were evaluated to observe how participants’ use of the tool developed throughout the semester. This included tracking changes in how the tool was applied, as well as recognizing links between reported use of the parallel corpus tool and language development. This analysis sought to identify patterns in how students used the tool, and strengths and weaknesses of the tool as a pedagogical instrument in language development. Again, the findings of this analysis were then compared with the findings of the other data sources.
In order to analyze the instructor’s pedagogical implementation of the parallel corpus tool, the instructor’s autoethnography was reviewed and analyzed. The analysis of the ethnography brought to light patterns of use, specific adaptations made by the instructor that improved the effectiveness of the tool, as well as identifying several key struggles and drawbacks identified during the study. The analysis of the autoethnography was also then compared with other data sources to consider the effect that the decisions and actions of the instructor had on the students’ language development.

Finally, learner attitudes and perceptions were analyzed using the language logs and the end-of-semester questionnaire. The views and experiences recorded in student reflections in the learner logs were evaluated and coded according to recurring opinions, findings, and reflections. These recurring themes, as well as several striking individual reflections were compared to describe student opinion at different stages of the study. This analysis was used not only to consider participant experience, but also to draw conclusions of how the corpus tool could be more effectively adapted in the future. The analysis of the end of term questionnaire was also juxtaposed to observations made in the learner logs. The questionnaire provided a more detailed inquisition that gave further insight into participant experience and opinion.

*Sections of this chapter also appear in Bluemel, 2014. Permission was obtained from the copyright holder of the journal publication to include this material, and permission details are included in the Appendix, Dimension 2014 Copyright Permission.*
CHAPTER 4
DATA ANALYSIS

As previously clarified in section 3.3. Data Analysis, five sources of data were collected and analyzed for the study: course material, screen recordings, learner logs, end of semester questionnaire, and instructor autoethnography. These data sources are categorized and analyzed in detail below. The course material and screen recordings are discussed under the same subheading, as the analysis of these two data sources was very integrated. Course material was analyzed longitudinally to observe how participants progressed and developed throughout the semester, and the screen recordings associated with exams and select assignments functioned to provide further insight into learner development. The learner logs and end of semester questionnaire are also combined into one section, as all these materials are participant reflections and analyses of their learning experience throughout the semester.

The 12 participants in the study received the same treatment, but came from two different classes. Table 4.2 includes a list of the 12 participants, using pseudonyms. The first nine participants were in Group A, and the final three participants listed (Jack, Kate, and Leo) belonged to Group B. As all individuals received the same treatment, all 12 participants’ data were analyzed individually. From these individual analyses, three primary categories emerged based on participant performance. This categorization according to participant performance takes into account participants’ grades and language development as well as their application of the parallel corpus tool. Table 4.1 depicts these three categories and the participants belonging to each category. Category 1, poor performers, includes four students who did not fully participate in the course. The course offering was a pass/fail course for which participants received elective credit. Individuals
were required to earn a cumulative grade of 70% to pass the course. Despite regular attendance, these four individuals rarely submitted assignments, were generally underprepared for exams, avoided active participation in the class, and did not apply the parallel corpus tool as instructed in their learning. Category 2, resistant to parallel corpus, includes individuals who performed well in the class, but were resistant to applying the parallel corpus in their learning. More specifically, individuals in this category typically used the parallel corpus as a resource to complete assignments and exams, but not to develop the language. Category 3, integrated parallel corpus, differed from the second category in how they applied the parallel corpus tool. These individuals used the parallel corpus to understand and develop the language, and not just to find answers. Most individuals fit quite clearly into one of the three categories, though there was some overlap as is discussed in section 5.1. Discussion of Learner Achievement. Cathy is the only participant who is listed in two categories, as her performance was at times more consistent with those in category 2 and at other times with those in the third category.

The emergence of these categories will be made more distinct in the following analyses. Section one of the analysis is organized according to the three categories of participant performance. The general findings observed in both the course materials and screen recordings are therefore presented under these subheadings. The final two sections are approached more generally and all materials are analyzed collectively. Chapter 5 then goes on to provide a discussion of all of these analyses, and how they inform and interact.
Table 4.1.
*

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Participants in this Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor Performers</td>
<td>Individuals who underperformed in the class and did not apply the parallel corpus as instructed</td>
<td>Adam, Brian, Frank, Jack</td>
</tr>
<tr>
<td>Resistant to Parallel Corpus</td>
<td>Individuals who performed well in the class, but did not apply the parallel corpus as instructed.</td>
<td>Hanna, Isabel, Kate, Cathy*</td>
</tr>
<tr>
<td>Integrated Parallel Corpus</td>
<td>Individuals who both performed well in the class, AND applied the parallel corpus as instructed.</td>
<td>Diane, Ellen, Gail, Leo, Cathy*</td>
</tr>
</tbody>
</table>

*Cathy is listed in two categories as her performance was more consistent with the integrated parallel corpus category despite her resistance to the corpus.

4.1. Course Material and Screen Recordings Analysis

Analysis of course materials and screen recordings resulted in the three categorizations of participants based upon performance and application of the parallel corpus tool. It is important to emphasize that while this categorization is functional for analytical purposes, the focus of the analysis is on discovering evidence of participant development, and if and how the parallel corpus productively mediated learner development. Accordingly, the analysis seeks out evidence of participants learning and internalizing lexical, grammatical and other functional linguistic constructs through the use of the corpus tool.

All assignments and exams were graded, and participants’ performance on these tasks became a preliminary measure of learner achievement. Table 4.2 provides a list of all participants, the percentage grade they received on individual assignments and exams, and then a cumulative grade. While the information in this table indicates performance
generally, the primary focus of the qualitative analysis then explores how participants apply their knowledge of the language and learned linguistic features in order to complete the assigned tasks. The analysis continues categorically as well as longitudinally considering participant performance during each of the three month-long sections of the study, looking first at student performance and production on assignments and exams and then cross-referencing data from screen recordings for evidence of development.

Table 4.2.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Going to China</th>
<th>Lesson 13</th>
<th>Lesson 14</th>
<th>Exam 1</th>
<th>Zodiac Story</th>
<th>Lesson 15</th>
<th>Lesson 16</th>
<th>Exam 2</th>
<th>Webquest</th>
<th>Lesson 17</th>
<th>Lesson 18</th>
<th>Review</th>
<th>Exam 3</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Adam</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 – Brian</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>30</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>20</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 – Cathy</td>
<td>100</td>
<td>96</td>
<td>100</td>
<td>98</td>
<td>100</td>
<td>100</td>
<td>96</td>
<td>95</td>
<td>100</td>
<td>98</td>
<td>88</td>
<td>98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 – Diane</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>96</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>98</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 – Ellen</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>94</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>98</td>
<td>99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 – Frank</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>18</td>
<td>55</td>
<td>20</td>
<td>35</td>
<td>44</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>44</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>7 – Gail</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>98</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>96</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>8 – Hanna</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>92</td>
<td>100</td>
<td>100</td>
<td>92</td>
<td>82</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>9 – Isabel</td>
<td>100</td>
<td>97</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>100</td>
<td>84</td>
<td>78</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>80</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>10 – Jack</td>
<td>0</td>
<td>76</td>
<td>40</td>
<td>91</td>
<td>100</td>
<td>90</td>
<td>96</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>82</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>11 – Kate</td>
<td>100</td>
<td>100</td>
<td>82</td>
<td>96</td>
<td>100</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>100</td>
<td>98</td>
<td>94</td>
<td>100</td>
<td>88</td>
<td>96</td>
</tr>
<tr>
<td>12 – Leo</td>
<td>100</td>
<td>97</td>
<td>90</td>
<td>94</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.1.1. Poor Performers

As indicated in Table 4.2, the four participants belonging to the poor performers category (Adam, Brian, Frank, and Jack) failed to submit several assignments during the study. Other than Jack’s performance during the second month of the study, these four students generally scored low on the assignments they did submit as well as on exams. Due to the many missing assignments that were not turned in, the extent of analysis of
their development is somewhat limited. Fortunately, all of these participants did complete all exams, and the materials that were submitted are analyzed below.

4.1.1.1. Month 1

During the first month of the study Adam did not submit any of the assignments, and the one assignment that both Brian and Frank did submit provides little detail for analysis as both individuals only partially completed the task by writing the new vocabulary items, but not composing sentences. Jack did the same on the Lesson 14 assignment, but he did attempt a few sentences on the Lesson 13 assignment. The sentences were basic in nature, though, and simply added a subject to the vocabulary items. This approach worked for two of the vocabulary items 多大 (how much/how old) and 开车 (drive a vehicle) – to successfully form the basic phrases: 你多大? (how old are you?) and 我开车 (I drive), but his attempt to form other sentences with vocabulary such as 岁 (years of age) suggests that he was either simply guessing or did not understand the material. His sentence 我岁 (I years), as well as several other responses that were composed in the same manner, are incomplete both structurally and functionally. Based on these few examples it is impossible to derive whether he struggled with the material or simply had not applied the necessary effort to successfully complete the task.

Jack’s performance on the exam, however, suggests that his low performance on the written assignment was due to lack of effort, as he did well on the exam. In comparison to other participants, the sentences he formed on the exam were basic, but grammatically and functionally correct nonetheless. For example, the last section of the exam provided pictures of several different animals, and asked students to “write a
sentence describing each picture”. While other participants who completed this section came up with a variety of sentences, Jack used the same sentence for each image: 这是一只 + name of animal (this is a + classifier + name of animal). Looking at the screen recording, Jack did use the website nciku.com in completing his exam to look up names and classifiers for three of the six animals listed in completing sentences. His use of the website and his sentences suggest that he was aware of how to construct the sentences (demonstrative + verb + numeral + classifier + animal), but had not completely mastered the vocabulary. He also used nciku to look up several characters in the open-ended question section of the exam. He used the drawing feature of the website to draw the characters he did not recognize to then look them up on the dictionary. Using this feature he was able to adequately respond to the questions. The mistakes he did make on the exam were on identifying the correct pinyin and tone of several vocabulary items.

In contrast, the other three participants only completed the first section of the exam, which required students to complete a chart filling the missing character, pinyin, and/or English of a term (see Figure 4.1).

Figure 4.1.
Example of Vocabulary Section of Exam.

Directions: Fill in blanks with the missing information (either characters, pinyin, and/or English meaning):

当然  __________  __________  
___________  __________  dog
它  ta  __________
漂亮  __________  __________

After reviewing screen recordings, it became evident that Adam, Brian, and Frank all used the nciku.com website in the same way to individually look up these items based on
the information provided. Additionally, it was observed that they did so gradually during the exam – looking up a term every several minutes. None of them made any attempt at the written portions of the exam.

4.1.1.2. Month 2

Jack performed quite well during the second month of the study, and analysis of his performance contrasts with the other three. Adam, Brian, and Frank’s effort during the second month of the study is very similar to month 1. Of the assignments that they did complete, the same approach was taken in that the basic vocabulary section was completed, but no attempt was made at writing sentences.

Frank submitted the zodiac story assignment, which was also screen recorded. On this assignment Frank only completed the section that asked students to identify the names of the 12 zodiac animals in Chinese and English. The screen recording revealed that Frank did use the parallel corpus tool to complete this portion of the assignment. On his screen-recording all three language formats (characters, pinyin, and English) were always opened, and based on the position of his cursor it appears as though he read through the English to identify animals and then navigate the corresponding Chinese text to find the character. While this approach is reverse from how participants were instructed to use the tool, it did still require Frank to analyze the text to some degree as word-to-word equivalents are not highlighted in complete texts on the parallel corpus. In other words, he could easily use the English to find the Chinese sentence an animal was mentioned in, but then had to work through the meaning of the Chinese sentence in order to identify which character(s) referred to that animal. While this observation is not equivalent to language development or internalization, it does demonstrate that the
parallel corpus required Frank to engage with the meaning of the Chinese text, in order to complete this task – which he did successfully.

Adam, Brian, and Frank’s performance on the exam was also similar to month 1, with the exception that all three individuals also partially completed the dictation portion of the exam. While their written performance is somewhat unremarkable in comparison to the first exam, one observation of note is how Adam used the parallel corpus in completing the task. All three individuals are observed using nciku.com similarly to how they did on the first exam to complete the vocabulary section. Adam, however, was observed using the parallel corpus to search for items in which only the pinyin was listed. He first searched for the term using nciku, which does not allow the user to indicate tone mark. Using nciku to search for the term 岁 (sui), which in English refers to a person’s years of age, 47 possible results were returned and most had different tones. He then used the parallel corpus, as it requires the user to enter the correct tone when searching using pinyin. Figure 4.2 is a screen shot of Adam inputting the term into the parallel corpus. The search results led to him quickly identifying the correct character. This observation does not suggest language development, but again it demonstrates the parallel corpus tool’s potential for more efficiently searching a term and working through a developmental task.

**Figure 4.2.**

*Adam’s Parallel Corpus Pinyin Search.*
Jack’s performance during the second month of the study contrasts both with the other individuals in the poor performance group as well as with Jack himself during the rest of the study. Jack successfully completed all assignments and performed extremely well on the exam. Consistent with the first month, the sentences he composed were basic in comparison to other classmates, but most were correct and met the requirements of the respective tasks. This was also true of Jack’s zodiac story assignment, which he received full credit on. Analyzing his screen recording revealed that Jack used the parallel corpus tool almost exactly the same as Frank did. All three language formats were always opened, and the cursor position indicates that English was the format focused on the most. While Jack’s use of the tool is similar, he completed the entire assignment – though there is insufficient data to conclude the extent to which the parallel corpus was used to complete the written portion of the assignment.

Jack also did well on the exam, again composing sentences that were basic, but that fulfilled the requirements of the task. Reviewing his screen recording makes it apparent that he relied on the parallel corpus to complete several writing prompts. For example, one question asked students to write where their family were from (你家人从哪里来?). Jack is observed searching the phrase “从哪里” (from where), and then identifying responses where the individual was from Japan (我从日本来). So Jack wrote, “我家人从日本来” (My family is from Japan). Jack’s search and response suggest some understanding as after finding an example sentence from his search he is able to then slightly restructure that sentence with the correct subject “his family” to appropriately answer the question. Jack is observed using the corpus in this manner for other questions
as well, indicating his ability to use the parallel corpus to complete the task, but not necessarily to develop his language ability.

4.1.1.3. Month 3

The final month of the study found little change in approach by the four young men in the poor performers category. The approach and performance on the few handouts that were submitted were consistent with previous months in that only the vocabulary sections were attempted, and only with minimal effort. The same was true on the third exam; however, the analysis of screen recordings did reveal some slight alteration. Adam and Brian primarily used the drawing feature on nciku.com for the queries they made to complete the vocabulary section of the exam, turning to the parallel corpus only to search using pinyin so tone mark could be indicated. Frank was a bit more direct in searching strictly for answers. He immediately opened the parallel corpus and displayed the content from Lesson 17 and 18 in all three language formats. The screen scrolls up and down several times suggesting he was using the three categories to more efficiently identify vocabulary items in the three formats.

Jack was also consistent with his prior use of the parallel corpus to search out answers and items for the exam. One further contrast between Jack and the other three that became transparent in this final analysis was Jack’s greater reliance on the parallel corpus and rarely using nciku.com. The other three tended to primarily reference nciku.com, whereas Jack rarely visited the site after the parallel corpus was introduced. One likely explanation for this was that Jack accessed the tool from an iPad that has a built in keyboard input that allows the user to draw the character. Figure 4.3 displays a screen shot of Jack searching for a term in the parallel corpus using this input method.
Figure 4.3.
Jack’s Parallel Corpus Search Using iPad Drawing Input.

The extent to which the others used neiku.com for this feature, and Jack’s transition to using primarily the parallel corpus on the iPad, suggest the students preference for this particular feature.

The observations made of Adam, Brian, Frank, and Jack demonstrate minimal language development and progression during the semester. Although the individuals evidently evolved in the use of their resources in completing tasks on some assignments and the exams, the primary focus of their efforts were to produce answers and not develop language ability. In the case of these four individuals, the application of the parallel corpus was not a transformational activity and did not typically function as a mediational tool in Chinese development. However, the analysis does suggest that these participants found the corpus to be a useful search tool and resource in completing several vocabulary and a few sentence composition tasks. Furthermore, several of the observations of how the parallel corpus was used by these participants (such as the pinyin
search feature and compulsory sentence evaluation when comparing language forms) highlight specific aspects of how it could be of great value in language development if applied appropriately.

4.1.2. Resistant to Parallel Corpus

The individuals who were placed in this category ultimately did quite well in the course. Cathy, Isabel, and Kate all earned final scores that were in the upper 90th percentile, and Hanna’s cumulative grade on the materials completed during the study was 82%. All of these percentages were passing grades, and these four participants fulfilled the necessary requirements to successfully meet the objectives of the course. The aspect of their performance that separated them from the third group was their application of the parallel corpus. The following analysis reveals how they used the corpus to query vocabulary items or search for specific items, but there was not strong evidence of consistent use of the parallel corpus as a mediational tool for developing conceptual knowledge of the language. This is a general observation, and there are a few noted exceptions where some of the participants in this category did use the tool in a way that evidenced development of specific lexical items or grammatical constructs. Isabel started off using the parallel corpus simply as a resource for searching answers, but by the end of the semester transitioned her usage of the tool to align more closely with participants in the final category, and she evidenced a great transformation in her performance and language development. Cathy also used the tool on several occasions as a mediational learning resource, but her inconsistent application led to her split categorization between this group and the final group of students.
4.1.2.1. Month 1

The participants’ use of the parallel corpus tool is not evidenced during the first month of the study, as they had not yet been introduced to it. However, several of the characteristics observed during this timeframe are consistent with how they began to use and incorporate the parallel corpus once they did have access to it. Understanding their approach to learning before gaining access to the parallel corpus informs both how they then chose to apply the tool (or not), as well as prompting further questions for a discussion of how the parallel corpus, or other technologies, might better be integrated to support a wide range of learners.

During this first month of the study, Hanna was the only participant in this grouping who underperformed, simply because she failed to submit two of the three assignments. She did perform well on the assignment she submitted as well as on the exam. Analyzing the work she submitted reveals a pattern of approaching learning where Hanna sought to find an answer and not to understand the material. Though the assignment she submitted fulfilled the requirements, it appeared somewhat rushed and all answers were very basic. This same pattern was observed on her exam, but with more data to analyze. In illustration, all of Hanna’s sentences on the final portion of the exam were only either three or four characters in length and were constructed either as subject + verb object (马跑步 / the horse runs) or subject + predicate adjective (龙很大 / the dragon is very big). While simple, they did still meet the requirements of the task. Hanna’s performance on this task is actually somewhat similar to the approach that Jack took. Analysis of her screen recording provides further support to the observation that her main goal was to seek a fast answer and not to understand or create meaning. To
complete the six sentences she created on the final task, she first searched individual
terms on nciku, and then used the query results to form sentences. For instance, for the
sentence 马跑步 (the horse runs), she first queried horse and then wrote the character (马)
and then searched to run and wrote the first word returned in the search results (跑步).
While it is acceptable to search out terms, Hanna did not exhibit any evidence that the
results of her searches were evaluated. This is first made apparent by how quick her
searches progressed on the screen recording, spending only 15-20 seconds per search
before moving on. Her lack of scrutiny of search results is further evident in the sentence
that she missed points on. Online she quickly searched the term “tiger” and then “angry”
and wrote the first returned result for each item to compose the sentence “老虎发炎”
(tiger irritated/inflamed). The term 发炎 (fāyán) refers to irritation or inflammation and
not to the meaning of anger Hanna was trying to express. The primary concept to
emphasize in this observation is that Hanna appeared to focus on finding answers, and
not on understanding and learning the language constructs and concepts or creating
meaning through the language.

Kate’s learning experience appears to be nearly identical to Hanna’s, though Kate
did complete all assignments during this portion of the study and throughout the semester.
Similar to Hanna, the sentences Kate constructed were all very basic, and her screen
recordings revealed the same approach to constructing sentences. One step further is that
Kate was also very open about her pursuit of the right answer and not the concept, as the
instructor noted several occasions when Kate blatantly requested an answer and would
grow frustrated when she received an explanation of a lexical item or grammatical
construct instead of being provided with a direct answer.
Isabel’s written performance is similar to both Kate and Hanna’s in many respects as she also constructed short and basic sentences that fulfilled the grade requirement but did not necessarily evidence mastery or conceptual development of the lexical item or grammatical construct. Isabel’s sentence composition followed the same pattern as both Kate and Hanna’s not just on the exam but also in writing sentences for her Lesson 13 and 14 assignments. She completed the tasks adequately, and received high marks on both of them, but sentences are all very simple being only three or four characters in length. On one assignment she composed sentences such as 它蓝色 (it is blue) and 狗叫 Bailey (the dog is named Bailey). While Isabel’s sentences appeared to be a repeat of the others’ performances, the screen recording analysis demonstrates that Isabel used online resources quite differently as she did not simply search for terms using English and then quickly transcribe the resulting Chinese search result like the other two. She did query a few English terms such as “to drive” and “black” and then used these words in her sentences. However, she also constructed several basic sentences using vocabulary and structures she was already familiar with. For example she wrote 它叫马 (It is called a horse) and 猫很小 (the cat is very small). Isabel was familiar with all of these characters from prior lessons, and used them to form simple, but satisfactory sentences. Additionally, none of these terms were searched online, again suggesting Isabel either knew them or referenced her personal notes or book to form these sentences. While her responses on course materials are quite similar to Kate and Hanna’s during the first month of the study, analysis does not suggest that Isabel was attempting to only find an answer. The above simple sentences, and the minimal use of online resources suggest that Isabel used additional materials as learning resources when her own knowledge was
insufficient, but typically first attempted to use her existing language skill to complete the task.

Cathy’s performance during the first month of the study contrasts starkly with the other three in this category. Cathy was one of the top students, completed all assignments exemplary, and performed well on the exams. On the assignments she included pinyin for all sentences even though this was not asked for. She did make some mistakes, but she also attempted to write more complex sentences. For example, she wrote the sentence 你的漂亮叫什么名字?* (What is your beautiful named?*) for the new vocabulary word 漂亮 (beautiful). The sentence is incorrect as the object that pretty is meant to modify is missing, but it does demonstrate Cathy’s attempt to use this new vocabulary item with the previously learned question structure for asking for someone’s name.

She later evidenced that she developed a functional understanding of this and other new terms on the exam. On the last task of the exam Cathy used the vocabulary item correctly twice. She wrote the sentences 老虎很漂亮 (the tiger is very beautiful) and 我漂亮的马叫 Larry (My beautiful horse is named Larry). The first sentence correctly uses the term as a predicative adjective linked by the degree verb 很 (hěn) instead of a copula as would be done in English. The second sentence uses the term as an attributive adjective correctly positioning the associative particle 的 after the adjective, but before the noun it modifies. These two sentences demonstrate that Cathy learned from the error she made on the assignment, and had internalized the meaning and function of the lexical item 漂亮 (beautiful) so that she was able to use it correctly as both an attributive and predicative adjective. Analysis of her screen recording revealed
minimal use of online resources, as she only used the website nciku.com once to draw and look up the term 动物 (animal), which appears in one of the exam questions. Overall, the analysis of Cathy’s course material and screen recording strongly suggest that she had learned and internalized several of the terms and concepts introduced during this portion of the study, and that she was able to produce her own meaning using these constructs. During this first month, there are far more similarities between Cathy and the participants in the third category (Diane, Ellen, Gail, and Leo), but her use of the parallel corpus during the second and third months of the study is where those similarities begin to diverge.

4.1.2.2. Month 2

The first assignment during the second month of the study was the zodiac story assignment, which required students to use the parallel corpus in order to first read the story and then to complete the written assignment. Cathy used the tool as directed in order to complete the task and her performance on this task is actually quite similar to most of the students both in this resistant to parallel corpus category and those in the integrated parallel corpus category. While the poor performers were observed to have all three language formats open and accessible while the task was completed, and cursor position focused primarily on English, the majority of participants in the latter two categories chiefly had only the Chinese and English opened and cursor position was generally focused on the interactive Chinese text.

Cathy is a prime example of this, as her cursor hovered above characters indicating where she was in the story, and she clicked on characters to see tone mark, and/or pinyin as needed. Figures 4.4 and 4.5, which are screenshots from Cathy’s
recording, illustrate this observed use of the parallel corpus. In both figures the cursor can be observed over one word, which is then highlighted in gray indicating where Cathy is in her reading. Figure 4.4 shows that Cathy clicked on the word to see tone marks, and Figure 4.5 reveals that Cathy clicked on this word a second time to view pinyin. This is how she, and the other participants mentioned, navigated through the story – using the parallel corpus as intended to provide additional elements of meaning only as needed to assist the learner in navigating the story and the meaning.

**Figure 4.4.**
*Cathy’s Use of Interactive Character + Tone Mark Feature.*

![Figure 4.4: Cathy's Use of Interactive Character + Tone Mark Feature](image1)

**Figure 4.5.**
*Cathy’s Use of Interactive Pinyin Feature.*

![Figure 4.5: Cathy's Use of Interactive Pinyin Feature](image2)

In examining her written assignment, it became apparent that she struggled with the second activity as she had written, crossed out, and then rewritten several characters.
Eventually she did complete the task correctly, which asked participants to identify characters for the 12 zodiac animals. This task required participants to not just write the characters for these animals, but also to comprehend elements of the story, as there are thirteen animals mentioned, but only 12 that are the zodiac animals.

Cathy’s other assignments were well done, but not with as much effort in sentence composition as noted during the first month. Most sentences were basic in nature such as 我从 Pennsylvania 来 (I come from Pennsylvania) and 欢迎, Bob! (Welcome, Bob!), and were not as dynamically developed as her typical work. The responses on her exam were better developed and more consistent with most of her other work. The final task on the exam asked participants to describe where they lived and where they would like to live. Her sentences were well thought out and clearly described her desire to visit and live with her grandpa for a time at his home in Norway. Analyzing her screen recording did not provide much further detail for analysis as she did not reference the parallel corpus at all during the exam and only used nciku to search for three words. She searched in English to find the Chinese characters for Norway (挪威) and grandpa (爷爷), which she used in the sentences she wrote. She also searched in Chinese for the term 美国 (USA) likely in order to verify that she had the correct character. Figure 4.6 is a screen shot of her search for this term that illustrates how pinyin is used to input characters, indicating that Cathy knew the pinyin for USA (Měiguó), but not necessarily the characters.

Figure 4.6.

*Cathy’s Search in Chinese using nciku.*
After selecting the first character option, she then searched the term to observe it was the correct character, before then using it on her exam. This demonstrates how Cathy used her resources to either check or develop her understanding, but not to substitute it.

Hanna’s performance on the zodiac story assignment is very similar to Cathy’s both in how she was observed using the parallel corpus and in her written responses. She also correctly recognized all of the animals, but she did make the mistake of adding the character 年 (year) after six of the 12 animals. For example, instead of just listing the character 牛 (ox) to identify the ox she wrote 牛年 (year of the ox). Both words appear in the text, but Hanna clearly did not draw a distinction between the two. In her other assignments, she exhibited the same approach as in the previous month and composed very simple sentences such as 欢迎你 (Welcome ‘you’) and 他马上到 (He immediately arrived), both of which are identical to expressions that appear in the textbook. Her approach to the final exam also remained unchanged from month 1 of the study. She did not reference the parallel corpus at all, and used both nciku.com and Google translate to look up single items in order to compose simple sentences. While she did not make any errors this time in her sentence composition on the exam, her sentences were all very basic, and many of the characters and words she used were first searched and found online.

Kate’s performance on the zodiac assignment is very similar to both Cathy’s and Hanna’s. The fact that she used an iPad to complete the zodiac assignment, though, allows us to see that while she, like the other two, had both Chinese and English language formats opened, she focused primarily on the Chinese. This is apparent as she zooms the screen in as seen in Figure 4.7 and only scrolls over to the English occasionally to
reference different aspects of the story. While Figure 4.7 closely resembles Cathy’s use of the corpus in Figure 4.4, it illustrates how the user zoomed in on the Chinese text, and then was observed focusing primarily on this text throughout the duration of the activity.

**Figure 4.7.**

*Kate’s Use of the Parallel Corpus Using an iPad.*

Though her other two assignments were rushed and produced quite basic work similar to her month 1 efforts, Kate’s exam was observed to be quite complex compared to her previous work. In response to the prompt asking students to write about where they would like to live, Kate wrote, “我非常想跟我姐姐一起住在中国 (*I really want to live together with my sister in China*)”. This sentence and others are notably more complex than previously observed from Kate, and her screen recording revealed that she used the parallel corpus in developing these sentences. For the terms 非常(*extremely*) and 一起(*together*) she searched them using Chinese and then presumably considered how to use them. This can be surmised as she used the Chinese input method (see Figure 4.8) to search for terms, demonstrating that she already knew the words and could recognize the characters, and then viewed the returned concordances (see Figure 4.9) to
observe how these characters were used in sentences. The two figures are queries of two separate terms that were both incorporated in the complex sentence she used on the exam.

Figure 4.8.
Kate’s Query Using Chinese Input.

Figure 4.9.
Kate Interacting with Returned Search Results.

She did quite well in clearly expressing her meaning in the above sentence as well as others. In this exam she noticeably evidenced application of the parallel corpus as a resource to help conceptualize the function of several terms in order to then express her own meaning and ideas in the sentences she composed. This differed from what had previously been observed in Kate’s approach to learning, and resulted in the construction of more complex sentences articulating her ideas, and not simply quick answers to fulfill
the task. While this observation was encouraging, it proved to be a singular event and inconsistent with her work before and after this.

The final participant in this category, Isabel, appeared to use the parallel corpus more similar to the individuals in the poor performer category as she completed the zodiac story assignment. She tended to have all three language formats opened simultaneously and scrolled over the English and pinyin most frequently. However, she did manage to correctly identify all the zodiac animals and perform well on the assignment. Her responses on other assignments were again quite basic and she failed to complete much of the sentence composition task for Lesson 16. This same performance was carried over to the exam where she notably struggled to complete the written portion, making several mistakes in her writing and failing to complete the last task.

The analysis of her screen recording together with the exam strongly suggest that she was having difficulty both in understanding the tasks as well as in completing them. In illustration of this point, one of the exam questions asked, “你家人从哪里来？ (Where does your family come from)”? As seen in Figure 4.10, Isabel searched for the term 哪里 (where) in the Lesson 15 text within the parallel corpus and viewed both the Chinese and English formats.

**Figure 4.10.**
*Isabel Searching Lesson 15 in the Parallel Corpus During the Exam.*
As observed in the figure, the text for the question is similar to exam prompt, except it has a different subject and asks ‘where are you from’ (你) and not ‘where is your family from’ (你家人). Isabel answered the question in the text and not the one from the exam, notably patterning her answer after the textual example as she wrote “我从日德国来 (I come from Jap-/Germany)”. First, the subject was different as she responded with I instead of my family, but then she also includes the first character for the word Japan (日), as this is the country listed in the textual example. From this analysis it becomes apparent that Isabel was attempting to use the parallel corpus as instructed as a mediational tool to help develop her own meaning creation in sentence composition. Despite this attempt, her efforts fell short, leaving the question as to whether this was a shortcoming of the functionality of the parallel corpus or Isabel’s understanding and ability to navigate it. At any rate, the corpus tool did not function to appropriately mediate and assist Isabel in her sentence construction and language development for this question or on the exam.

4.1.2.3. Month 3

The final month of the study marked improvement in Cathy’s development and evidenced that she had internalized several of the lexical items and linguistic constructs that had been learned. The Lesson 17 assignment was quite illustrative of this development as it asked participants to use the parallel corpus to read a letter from Xiūmíng about his family, and then write back to him. Cathy demonstrated the correct use of new location phrases that were first encountered through the text. She drew a picture of her family and then described the picture with sentences such as 我在最右边 (I
am on the far right) and 我旁边是我妈妈 (My mom is next to me). Both 右边 (right ‘side’) and 旁边 (next to) were new terms, but Cathy was able to individually identify their meaning and correctly compose the sentences. This is particularly worth noting in the second example sentence, as the word order in the Chinese expression is quite different from in English. The sentence might more literally be translated as “My next to is my mom,” and many of the participants struggled to create grammatically coherent sentences using the term 旁边 (next to), but Cathy was able to do so without problem.

Evidence of her development and functional understanding of these terms is apparent on the final exam when she again uses these same concepts in constructing similar sentences describing the occupations of family members. While there is evidence of development of this and other language concepts in Cathy’s course material analysis, her feedback and screen recording analysis are not able to link this development with her use of the parallel corpus. For the final exam the only online resource she used actually was the parallel corpus, but she used it to search in English for a few specific characters such as 朋友 (friend) and 弟弟 (younger brother), and primarily to view the complete text associated with Lesson 17 with all three language formats (Chinese/pinyin/English) displayed. She had this text open for about 15 minutes and is observed scrolling back and forth several times with her cursor rolling over the English and then across to corresponding Chinese and pinyin. Her use of the tool to view all language formats focusing first on English suggests that when she did use the parallel corpus it was to quickly search out an answer, and not to understand the concept. While this use of the tool is not necessarily surprising during the exam period, it does contrast with the
approach taken by the participants who integrated the tool in their language development, as is further discussed in section 4.1.3.3.

Hanna and Kate again performed very similarly on the written assignments, and both again used short sentences in composing sentences. This stood out on both of their Lesson 17 assignments in which they described their family to the hypothetical pen pal Xiūming. While Cathy and the participants in the integrated parallel corpus category tended to write longer sentences using some of the lexical and grammatical structures modeled in the text, Hanna and Kate stayed simple with sentences such as 这是我的家人 (This is my family), 这是我 (This is me), and 那是我的爸爸 (that is my dad). Consistent with what has been observed in the other work completed by these two, the answers they provided were very basic, but did serve to meet the requirements of the assignments. This same observation is again repeated on the final exam administered during the study. Both participants received scores in the 80th percent for this exam, primarily due to struggles with the final task on the exam, which required students to write a 40-character essay. Hanna and Kate wrote grammatically and functionally correct sentences, but they were again very basic and short. This led to both participants falling short of the required 40-character mark, writing 30 and 32 character essays, respectively. The struggle with this task appears to be largely due to lack of mastery of the concepts and vocabulary needed to complete it. In analyzing their screen recordings, it became apparent that neither student used the parallel corpus and Hanna mainly used nciku.com while Kate relied on Google translate. Both individuals used their respective source to repeatedly search in English for Chinese terms to use in their writing, but this approach left them
short on time and ability to complete the task, and more importantly evidenced that while they knew resources to turn to, they themselves had not internalized the material.

Turning to our last participant in this category, Isabel, we are able to observe both conceptual and functional development during this final month of the study. After receiving a score she was displeased with on the second exam, Isabel approached the instructor for further assistance. During this individual meeting with the instructor, many of the main concepts that had been covered were reviewed, and practiced. Additionally the instructor also discussed with Isabel potentially more efficient ways for using learning resources such as her textbook, online dictionaries, and of course, the parallel corpus. Following this meeting, her performance on Lesson 17 began to demonstrate improvement. Though the reported score for Isabel is 100% on this task, that score was only earned after submitting a modified second version of the assignment. Isabel attempted to develop more complex sentences in her initial submission of the assignment, which required her to write a letter describing her family. She made several mistakes with new terms when indicating direction, and wrote sentences such as 妈妈旁边爸爸* (mom next to dad*), and 我不姐姐* (I not older sister*). The instructor returned the assignment with errors marked and directed her to reference the parallel corpus and evaluate sentences using the same terms she struggled with. Isabel resubmitted the assignment with the sentences 妈妈旁边是我爸爸 (My dad is next to mom) and 我没有姐姐 (I don’t have an older sister(s)). Both sentences were patterned after examples found in the text, but express the correct meaning Isabel was trying to create.

Similar experiences and outcomes were noted on both the Lesson 18 assignment and the review assignment. Isabel made efforts to express more complex ideas in more
dynamic sentences, and the instructor allowed her to correct and resubmit both assignments after correcting mistakes similar to those noted in Lesson 17. Isabel’s performance on the third exam continued consistently with the development observed on her assignments during the final month of the study. While her exam score was lower than she desired at 84%, her improvement on the exam is very noticeable as she constructed more complex sentences, and did so well. Similar to Hanna and Kate, the points Isabel missed were primarily a result of her not being able to complete the last task on the exam. However, the sentences she did write were far more complex than what was observed in previous months, writing sentences such as 我两个妹妹没做工作, 和我没做工作 (My two younger sisters have never worked a job, and I have never worked a job). The sentence is notably longer, and far more developed than her previous materials. Although she misuses the conjunction, 和 (and), she does evidently demonstrate an increased effort and development as and draws upon the resources available to her to convey her meaning. Looking at her screen recording, it is also evidenced that the primary resource she used was the parallel corpus. She used nciku a couple times to search in Chinese for the terms 别的 (other) and 成员 (members), which both appeared in the prompt for the final task. Otherwise, she referenced only the parallel corpus and is observed viewing the letter from assignment 17 detailing Xiūmíng’s family and the text from Lesson 18 about occupations. In both cases she displayed the Chinese and English side by side, and is seen then scrolling over the Chinese similar to how participants used the tool in completing the zodiac story assignment. This use, and the resulting sentences Isabel wrote that appear to be patterned after textual examples suggest that Isabel was using the parallel corpus to first find example sentences and then to analyze in order to
conceptualize the ideas being expressed so that she could then produce her own meaning using the various constructs. This observation also suggests that by the end of the study period, Isabel was starting to use the parallel corpus as a mediational tool to help her understand and develop her language ability. Though she still had progress to make, there was noticeable improvement from previous months, and observed reliance on the corpus tool in aiding her development.

4.1.3. Integrated Parallel Corpus

As with the individuals in the previous category, all participants in this final category also performed well in the course during the duration of the study and all had cumulative scores that were over 90 percent. In addition to achieving good scores on exams and assignments, the analysis of Diane, Ellen, Gail, and Leo’s course material and screen recordings demonstrate that all of the individuals consistently incorporated the use of the parallel corpus in a way that evidently led to their language development and conceptual knowledge of Chinese lexical items, grammatical constructs and composition. As previously mentioned, Cathy also exhibited some of these same characteristics, though not as consistently as the other four. Their learning experiences are analyzed below in juxtaposition to the participants in the previous two categories.

4.1.3.1. Month 1

Similar to the observations made in the other two categories, the analysis of Diane, Ellen, Gail, and Leo during the first month of the study begins to provide insight into who they are as learners and how they approach assignments, exams, and language development tasks. While some of their approaches are sometimes similar to other
participants, the contrast between participants in all three categories begins to emerge in the first month of the study before the parallel corpus was ever introduced.

Diane was the top-performing participant during the first month of the study, and received perfect marks on all three assignments and the exam. Her first two assignments were meticulously completed, but use of vocabulary and grammatical structures evidenced basic applications that mimicked the example sentences found in the textbook, and were quite similar to what was observed in Hanna, Isabel, and Kate’s assignments. For example, one of the simple sentences she produced on the Lesson 13 assignment was 我十五岁 (I’m fifteen years old). The sentence is correct and fulfilled the task on the assignment, but it does appear exactly as it is written in the book, even though Diane herself was not 15 years old. In her learner log she also noted her primary reliance on the textbook in composing these sentences, suggesting that she used resources to find answers, but not to develop understanding of concepts. Her third assignment, Lesson 14, exhibited more original sentence composition as she described sentences about her personal pet dog. While sentences were still basic, they begin to demonstrate Diane’s ability to use the new grammar and lexical items to construct her own meaning. Several example sentences include: 我有一个狗 (I have a dog), 它是我的朋友 (S/he is my friend), and 它叫 Ru (S/he is named Ru). The first sentence does have an incorrect classifier (个 geì instead of 只 zhī), which the instructor marked on her assignment, but as students had not yet learned the classifiers for animals no points were subtracted. Of more importance is the observation that Diane had correctly identified the need for a classifier in this sentence and used the general classifier 个 showing her understanding of the need for a classifier, but that she did not necessarily know which one to use.
Participants had previously learned about classifiers and been introduced to a few of them, but the discussion of classifiers for animals occurred the day following when this assignment was submitted. The attempted use of a classifier, as well as the original meaning being expressed by Diane in these sentences suggests her developing understanding of both new lexical items and previously learned grammatical constructs.

By the exam Diane demonstrated her growing understanding of classifiers, and that she had learned which to use for animals. For the final task on the exam that required students to describe images of various animals, Diane was the only participant to compose all sentences in a way that required the use of a classifier, and she used the correct classifiers depending on the animal, and composed sentences such as: 这匹马是黑色的 (This horse is black) and 这只老虎美丽 (This tiger is pretty). This observation demonstrates that Diane exhibited development of the classifier construct. She evidently understood how classifiers functioned, as she had now learned which classifiers applied to each animal, and could correctly form her own meaning with the correct classifier. Examining her screen recording was somewhat unremarkable. She used both nciku.com and mandarintools.com, but only to look up two terms – the color black to compose the above sentence, and the vocabulary item 开车 (to drive), which appeared earlier in the exam.

The observations made in analyzing Gail’s performance are very similar to Diane, and Gail only had points marked off for a portion of one assignment where she misunderstood the directions and wrote pinyin instead of characters. Similar to Diane, her performance on the exam was also very good and sentences were all correct, but they
were basic in nature. Her screen recording revealed that she only used nciku.com once to look up the color term 灰色 (gray).

Ellen’s overall performance during the first month of the study was just as impressive as both Diane and Gail, but more clearly illustrated Ellen’s understanding of material both on assignments as well as the exam. Ellen tended to construct more dynamic and complex sentences than her classmates. She demonstrated that she sought to create her own meanings and not copy or imitate textual examples. For instance, she composed the sentence, 我的猫很漂亮和喜欢吃肉 (My cat is very beautiful and likes to eat meat), to use the new vocabulary item 漂亮 (beautiful). This sentence combines vocabulary and elements from several prior lessons and uses the conjunction 和 (and) with a second descriptive clause. While this sentence is not necessarily complex, for a beginning level student this is a longer sentence, and Ellen’s regular construction of more complex and dynamic sentences such as this one contrasts with her classmates.

The few minor mistakes Ellen made on the exam likely resulted from her desire to construct more complex sentences like the one described above. For example, the sentence she provided in response to the question 你有什么动物? (What animals do you have?) was 我有一只狗和猫两只* (I have a dog and cats two*). The sentence again uses the conjunction 和 (and), and its intended meaning is clear but misplaces the noun 猫 (cat), which should follow the classifier instead of coming before it. Thus while Ellen did miss a few minor points on the exam, the analysis actually reveals her understanding of the language and ability to produce meaning excelled beyond her classmates, and the complexity of her sentences should perhaps have been taken into more account in the grading process. The last note to make was that Ellen’s screen recording revealed
minimal use of online resources. She only referred to an online dictionary to look up characters for colors, which she then used to describe the color of animals in the final task.

Leo’s performance during the first month of the study is most similar to Ellen as he also attempted to construct more dynamic and purposeful sentences, but his skill wasn’t quite at the same level as Ellen. The primary observation to note from analyzing Leo’s material was how he used the online resources. He used nciku.com to search for unknown terms that we had not yet used in class in order to make his sentences more dynamic. Although Hanna and others were observed doing this same thing, Leo seemed to do so more analytically as he did not always select the first item returned in a search query, and he also further searched the items he chose to use in order to view example sentences that contained the word. For instance, he searched for the expression “every day” and while the first result listed was 日常 (rìcháng), after looking at several options and their example sentences he ended up using 每天 (měitiān). The sentence he composed for the exam to describe an image of grumpy cat was 猫每天很生气 (The cat is angry every day). This sentence correctly uses the term he searched for. While there were also some instances where he made mistakes in word order while attempting to compose more dynamic sentences, the analysis of his screen recording suggests he used the website not just to look up terms but also to evaluate how those terms were used and mimic the patterns.

4.1.3.2. Month 2

All the participants in this final category started the second month of the study off by performing very well on the zodiac story assignment. While Diane did a great job as
well, after reviewing her screen recording it became apparent that this wasn’t an easy task for her. Initially, she used the parallel corpus similar to how the participants in the poor performer category did in that she opened all three formats (Chinese, pinyin, and English), and appeared to be looking first at the English, then the pinyin, and then trying to identify characters that corresponded. This is apparent at one point when she is trying to search for the character associated with the pinyin 大帝. The pinyin was part of the expression 玉皇大帝 (yuánhángxàndì), which refers to the great jade emperor. After reading the English, Diane read through the pinyin and was unsure what this referred to so she tried to search for the character using the pinyin. She did this first using Google, then mandarintools.com, and then finally on the parallel corpus. She eventually found the characters, and figured out the meaning, but it was not without great effort. She followed this same process for a few other terms, and was about twenty minutes into the activity when she adjusted her approach. About twenty minutes in she ended up hiding the pinyin column in the text, and then working through the Chinese for the rest of the passage similar to Cathy, Hanna, and Kate – in that she began focusing on the Chinese and then only using the interactive feature to identify characters she was not certain of. From this point on, she moves more efficiently through the remainder of the text, and appears to have learned and experienced how to use the parallel corpus more effectively.

Her other two assignments are somewhat unremarkable. She did very well on both assignments, and created good sentences expressing individual meaning but there is not evidence of either struggle or use of resources in completing the tasks to comment on. On the second exam, though, there are some observations worth noting. Diane performed well on the exam, constructing sentences that evidenced understanding of the terms and
concepts she was using. What does stand out is how Diane used the parallel corpus similar to Adam in order to search for a term using pinyin so that she could narrow the search by indicating correct tone mark (see Figure 4.2). She even searched the same term as Adam, sui (岁, years of age), first using nciku, then mandarintools.com. Both resources returned numerous results, so Diane then turned to the parallel corpus and searched the pinyin indicating the tone and returned more narrow search results, and quickly moved onto to the search of another item indicating the corpus had helped her resolve this issue. This observance is worth noting for two reasons: first, as previously mentioned the participants found this feature of the tool useful as it functioned to refine their search. Second, this observation was unexpected as students typically struggle most with remembering and learning the tones. The fact that they relied on the tone to identify the character indicates both internalization of the tone and pronunciation, and also an unexpected benefit of the parallel corpus design.

Gail and Leo’s experience during this second month of the study is very similar to that of Diane, with a few important distinctions to highlight. First, while it took Diane about twenty minutes before she began to realize how the parallel corpus could function more efficiently if she focused on the characters instead of the English, Leo approached the assignment this way from the beginning. Gail was absent on the day this assignment was completed, and never submitted it so no comparison can be made with her. However, Leo’s performance was similar to Cathy, Hanna, and Kate as he had both the Chinese and English versions of the story opened, but appeared to focus primarily on characters. Leo’s performance on this activity as well as his and Gail’s performance on the other assignments was very good and quite comparable to Diane.
They all also did very well on the exam, but exhibited some variation in how they used the parallel corpus to accomplish similar tasks. While Diane had found it helpful on the exam to take advantage of the pinyin search feature that narrowed search results for characters, Gail appeared to find this same feature less than helpful when searching for a term that she did not know the pinyin for. Figure 4.11 shows Gail’s initial search in the parallel corpus for the term 开车 (to drive a car).

![Gail’s Unsuccessful Parallel Corpus Pinyin Search](image)

As the search results show, she had first searched the term selecting the neutral tone for all vowels, so the query returned no results. She tried searching a couple more times, guessing at the correct tone, but then changed her approach and used nciku.com for this search, and was able to identify the character she was looking for without having to indicate tone mark. She did come back to the parallel corpus to search the term again using the characters she had identified, but it was evident in this and a couple other instances that Gail struggled with remembering tone marks and therefore had a hard time taking advantage of the pinyin search feature in the corpus tool.

One additional aspect of Leo’s experience to point out is that similar to Jack, he rarely used nciku.com or resources other than the parallel corpus after being introduced to it. Jack and Leo were in the same group, and both had access to iPads, which have the
built in character drawing Chinese input system (see Figure 4.3). Having this feature greatly reduced Leo’s reliance on other resources, and he was only observed using the parallel corpus during the exam.

Ellen’s first experience in using the parallel corpus is slightly different than what was observed among all other participants, but precisely what the instructor had advised and hoped students would do. As the screen shot shown in Figure 4.12 illustrates, Ellen opened only the Chinese version of the text on her screen. It is very easy to follow where she is at in her reading, as she followed along with her cursor allowing word segmentation to be highlighted in gray as she scrolled over terms. In the image her cursor is over the term 知道 (to know). She primarily used only this version of the text and only on four occasions did she open the English version. After identifying the term or portion of meaning she was missing she would hide the English once again and focus on the Chinese. She is observed clicking on several different characters she was unfamiliar with to have the tone marks and/or pinyin highlighted, similar to what was observed in Cathy’s interactive use of the tool in Figures 4.4 and 4.5.

Figure 4.12.

Ellen’s use of the Parallel Corpus During the Zodiac Assignment.

Her performance on the task was very meticulous and accurate. One further note to make is that the last task of this assignment asked students to write a brief summary of the content of the story. This task has not previously been mentioned as participants
completed it using English. However, Ellen was the only individual to attempt her summary in Chinese. It is mixed, and begins with the sentence “中国的 emperor xihuān know 他几岁” (China’s emperor wanted to know how old he was), which is a mix of characters, pinyin, and some English. The primary observation to be made is the extra effort often observed in Ellen’s work and study efforts.

This extra effort was apparent on her work on other assignments, as well as on her exam. Ellen typically composed the most complex sentences, and repeatedly evidenced a greater level of aptitude and development of the language in comparison to her peers. Ellen’s purpose on activities and exams seemed to focus on developing her Chinese language knowledge to convey the meaning she wanted, rather than limiting herself to express only what she had already learned. In other words, she did not allow herself to be limited by her developing vocabulary and language knowledge. One further specific aspect of the analysis to note is how Ellen used the parallel corpus on the exam. Similar to how she approached the zodiac assignment, she primarily referenced the Chinese text, and used the interactive tone mark and pinyin feature only as needed. Figure 4.13 illustrates her referencing the Lesson 16 material highlighting the term 住 (to live).

**Figure 4.13.**

*Ellen’s use of the Parallel Corpus During Exam 2.*
This much information proved to be sufficient for the sentence she was working on and she did not need to explore further formats. This is how she was generally observed using the parallel corpus on the exam, only referencing the English format on three occasions. What stands out about this experience is that Ellen applied the corpus as a mediational tool to assist her development both on assignments as well as on exams – focusing on understanding of concepts and not on seeking out answers.

4.1.3.3. Month 3

Diane’s experience on assignments during the final month of the study is most noteworthy, as she reports using only the parallel corpus and notes as resources. Similar to Cathy, she formed quite complex sentences on assignment 17 by emulating the grammatical structures in the text in order to describe a picture of her family and their location in reference to one another. Her performance is actually very similar to Cathy’s on this task. A similar performance is observed in assignment 18, as she described the occupations of her family members with sentences such as 我爸爸妈妈都是大学老师 (Both my dad and mom are university teachers). Interesting to note in this sentence is that Diane correctly used the term 都 (all/both), even though the grammatical use and function of this term had not been covered in class – suggesting that she either learned the word from the parallel corpus or the textbook. Furthermore she used two terms she was familiar with, 大学 (college/university) and 老师 (teacher), to describe her parents’ occupation. There is a separate term that can be used for college professors (教授), but Diane was able to use the resources and knowledge that she already had to accurately convey her meaning.
Her performance on the exam duplicates much of the work observed in her assignments, and the speed with which she completed the exam as well as her lack of reliance on resources suggest that Diane was very comfortable with the new material by the time of the exam. The only resource Diane used for the exam was the parallel corpus, but she only referenced it once to look up the term big (大) to find the correct tone mark for the pinyin. Other than this one query Diane did not use any online resources, and she was the first individual to complete the exam. While participants in both the other categories struggled to complete this exam in the allotted time, Diane managed to do so with time to spare. As her use of the parallel corpus on the assignments is evident, it is quite possible the use of the parallel corpus helped Diane in developing conceptual and functional knowledge of the material covered during this term. Though this explanation is quite plausible, there is not enough evidence in the data to make this correlation firm.

The performance of both Gail and Leo during the final month of the study is also very similar to Diane. Their performances on the assignments were also well thought out and developed. At one point, in Gail’s Lesson 17 response letter introducing her family, she states her surname two different ways, demonstrating her understanding of the newly introduced term 姓 (surname) and how it could be used as a verb, or combined with the character 名 (name) to function as a noun. Her sentence reads, 我的姓名是 Smith，我姓 Smith (My surname is Smith, I’m surnamed Smith). This example and others again illustrate that Gail had a developed conceptual knowledge of terms and could use them accurately. Sentences similar to these two existed in the parallel corpus, and based on Gail’s self-report in her learner log of using the corpus in searching this term, it can be deduced that the corpus mediated her understanding of this and other new concepts.
Furthermore, Gail then went on to use this term again in the exam when describing her father and his occupation as a professor. His last name differed from hers (again, pseudonyms are used), so after introducing herself she explained, “我姓 Smith. 可是我爸爸的姓名是 Smithfield。他是大学老师。(I’m surnamed Smith, but my dad’s last name is Smithfield. He is a college professor). Her use of the term again both as a verb and a noun infers that she had internalized this concept, and her self-report of using the parallel corpus to first understand this term again suggests its role in her development. During the exam, Gail is witnessed only referencing the parallel corpus for a few minutes, and looking only at the Lesson 18 text, which was open with both the character and English formats visible. Her cursor is observed over vocabulary items for various occupations, indicating she primarily used the parallel corpus for occupation vocabulary. While Leo’s performance on assignments and the exam is quite similar to Diane and Gail in many respects, there is one behavior observed in his use of the parallel corpus during the exam that is worth drawing attention to. Figure 4.14 illustrates Leo’s search of the term 时候 (time/time period). This term had been encountered in a few previous supplementary texts, but not explicitly covered in the class. Leo’s search of the term using Chinese first indicates that he knows term, and likely its meaning. In the screen recording he goes on to scroll over the 12 results returned from the query, presumably analyzing usage. The presumption is confirmed in his exam essay when he writes, “我去去年圣诞节的时候想是护士。我现在想是医生。(Last year at Christmas I wanted to be a nurse. Now I want to be a doctor.)”. This 时候 (time/time period) construct functions differently in Chinese than in English and a more direct translation of Leo’s first sentence might read I last year Christmas at that time wanted to be a nurse. Leo managed to use it
correctly, but his reliance on the parallel corpus suggests he needed further mediation to review and correctly use the term in composing his sentence. He is observed searching a few other terms, including 现在 (now) that appears in his second sentence, and using the parallel corpus in the same manner. The only other use of the tool observed in his screen recording was to use the drawing feature on the iPad to look up a couple characters from the questions in the exam that he was unfamiliar with.

**Figure 4.14.**
*Leo’s Search Using the Parallel Corpus During the Exam.*

Ellen’s performance once again stood out from her peers during this final month of the study. On the Lesson 17 assignment, Ellen used several of the direction terms from the text in composing her writing – the same as many other students had, but she also went beyond that. She included additional sentences using several complex and challenging words. For instance, she added the introductory sentence “这是我全家福 (This is my family portrait)” . While the structure of this sentence isn’t necessarily complex, the word 全家福 (family portrait) is a very specific and uncommon vocabulary word that appeared once in the original text. In Ellen’s writing she not only used several
of the main direction terms and grammar structures as noted, but also included some of the more specific and complicated words. When discussing this assignment later in class, Ellen was also the only class member able to identify several of the terms. As the text for the Lesson 17 assignment was all on the parallel corpus, it is easier to conclude that the parallel corpus functioned successfully in Ellen’s development of the lexical items and grammatical constructs mentioned, and aided in her composition.

Lastly, as Figure 4.15 draws attention to, Ellen’s use of the parallel corpus on this exam was consistent with how she was observed using it both on the previous exam and on the zodiac story assignment.

Figure 4.15.

Ellen’s Use of the Parallel Corpus During Exam 3.

She opened the Lesson 18 text, and primarily referenced only the Chinese. In the figure she has both the English and Chinese open, but after about twenty seconds she hid the English and looked over the Chinese. Once again this use of the tool indicates that she was applying the parallel corpus as a mediational tool for her language development.
4.1.4. Final Presentation

Initially, the final presentations were not envisioned as part of the data analysis. On the second to last day of the class, participants all gave a brief final presentation on something they found intriguing about the Chinese language or culture. Each individual was required to use a PowerPoint for the presentation. Students were not required to complete any written task for this assignment, but they were asked to submit a copy of any written materials that they chose to use – whether that be handouts, notecards, or anything else. Five participants ended up submitting materials, all of which were scripts used during their presentations.

What immediately became striking about these texts is the format in which the participants had prepared them. Two individuals, Cathy and Diane, had both handwritten their scripts using only Chinese characters. This observation was encouraging, as they had not used pinyin at all for reference while giving their presentations. Students typically rely heavily on pinyin and are less confident with characters. The fact that both of these participants chose to write in only characters, suggests that they both had confidence in their understanding of the characters they had written, and their desire to focus on using characters and not pinyin.

The other three individuals that submitted materials (Ellen, Gail, and Isabel) also submitted scripts used for their presentations. Their scripts were different from the other two, as they were type written and all were prepared in three language formats: Chinese characters, pinyin, and English. Figure 4.16 depicts Ellen’s script, and is representative of how both her and Isabel prepared their scripts, with the language formats being aligned paragraph by paragraph. Significantly, both individuals listed the formats in the same
order: characters, pinyin, and then English. Gail also used the three-language format in preparing her script, but as seen in Figure 4.17 she aligned her script sentence-by-sentence, and positioned pinyin before characters.

Figure 4.16.

*Ellen’s Final Presentation Script.*

我的陈述是关于大熊猫。大熊猫住在竹林。它们主要吃竹子。它们也吃草和小动物。大熊猫花了大量的时间吃。它们不旅行一起，可是妈妈遇到它的幼仔。幼兽很小。大熊猫的大小美洲黑熊。大熊猫是濒危物种由于栖息地丧失和漂洗。它们是保护的标志。中国有保护区还禁止漂洗。中国有大约地四十大熊猫保护区。

wǒ de chéngshù shì guān yú dàxióngmāo。dàxióngmāo zhū zài zhúlín。 tāmen zhù yì chāo yǐ xiào。 tāmen yě chǐ cǎo hé xiǎo dòngwù。 dàxióngmāo huále dāliǎn de shūjīn chǐ。 tāmen bù lǚxíng yíqǐ, kěshì māma rào xǐ tāde yǒu zǐ。 yóushòu hěn xiǎo。 dàxióngmāo de dàxiāo Méizhōu hēixióng。 dàxióngmāo shì bǐnwéi wūzhōng yǒuyú qǐxī shīqū hé piào xī。 tāmen shì bāchú de biǎozhì。 zhōngguó yǒu bāo hú qū hé jǐnzhī piào xī。 zhōngguó yǒu dàyuè de sìshí dàxióngmāo bǎo hú qū。

My presentation is about giant pandas. Giant pandas live in the bamboo forests of central China. They mainly eat bamboo. They also eat other grasses and small animals. Giant pandas spend much of their time eating. Giant Pandas do not travel in groups, expect for mothers and cubs. A baby panda is very small. Giant pandas grow up to be about the size of an American black bear. Giant pandas are an endangered species because of habitat loss and poaching. They are a symbol of animal conservation. China has reserves and prohibits poaching. China has about forty reserves for giant pandas.

Figure 4.17.

*Gail’s Final Presentation Script.*

Zhōng guó de yǐnyuē yǒu xiǎn yuè qì。mù guān yuè qì。hé cǎ jī yuè qì。

Chinese music has strings, woodwinds, and percussion.

Yì xiě shòu huānyìng de yuè qì yǒu gǔqín, gǔzhēng, pīpā, érhú, hé dízì。

Yì xié 受欢迎的乐器有 古琴，古筝，琵琶，二胡 和 筝子。

Some instruments are guqin, guzheng, pipa, erhu, and dizi.

Wǒ bāba de péngyǒu biāoyǎn érhú。Zhè tǐng qǐlái zhēn hǎo。

我爸爸的朋友表演二胡。这听起来真好。

My dad’s friend plays the erhu. It sounds very pretty.
Ellen and Isabel’s formatting suggests an emphasis placed on characters, whereas Gail appeared to place emphasis on pinyin. What is quite remarkable about all three individuals is that they conceptualized their final project using the same type of textual alignment they interacted with on the parallel corpus tool. There were no requirements, suggestions, or other indications of how to create a script, and as it was a spoken presentation students could have chosen to use only the pinyin, as that would have been functional and easier.

The fact that Cathy and Diane included only Chinese infers their transition from concentrating on pinyin to focusing on characters. The three-language format of the other three strongly suggests that they had not only developed their language ability through the parallel corpus using these formats, but that in doing so they have begun to conceptualize their interaction with the written language through these formats. To the extent that this is positive or negative is yet to be explored, but it does strongly evidence the role of the parallel corpus as a transformative mediational tool in these participants language development.

4.3. Learner Log and Questionnaire Analysis

The initial reception of the new online parallel corpus tool was very positive. As previously indicated, the participants were familiar and proficient with a wide variety of technological resources and tools before the implementation of this study; and were generally encouraged by experience with new technologies. This was observed several times prior to the introduction of the parallel corpus tool, as students’ demonstrated similar attitudes of excitement when other online resources were introduced. Participant feedback and comments in their learning logs demonstrated how these resources
(nciku.com, readchinese.nflc.org, www.eon.com.hk/estroke/, and others) had been adapted and integrated into their learning experience. As the participants were proficient with a variety of language learning resources and technologies, their learner logs proved valuable in demonstrating not only how they used these tools, and which they preferred, but also in tracking what features the participants felt were lacking and that they would like to see in an ideal language learning tool.

As indicated on the learner log prompt (see Appendix, Learning Log Prompt), participants were asked to respond to six questions every time they completed an assignment or exam. These prompts asked them to: 1) list resources they used to complete the task, 2) describe which of those resources were most helpful and why, 3) discuss what was most challenging about the task, 4) consider what possible resources would make the task less challenging, 5) describe what was most beneficial for them in completing the task, and 6) provide any further thoughts about their learning experience. Though the parallel corpus tool is not listed explicitly in the learner log prompts, it was implicitly inferred as participants were asked to analyze all resources used and the corpus tool became an integrated learning resource during the course of the study. Accordingly, many of the learner logs explicitly analyze and provide feedback on the parallel corpus tool after it was introduced.

In total there were eight separate assignments and three exams for which the participants completed learner logs. Learner logs were not completed for the You’re Going to China! and Webquest assignments. All participants completed the exams, but not all individuals submitted each homework assignment. Additionally, several participants submitted homework without completing the learner log portion of the
assignment. Of the 132 possible learner log submissions between assignments and exams, 111 completed or partially completed learner logs were received. Participant responses on these learner logs, as well as responses to the end of semester questionnaire are presented in four sections below categorized sequentially according to when the students completed them. The first section includes the logs completed during the first month of the study before the parallel corpus tool was introduced (Lesson 13, Lesson 14, Exam 13 & 14). Section two discusses the logs completed during the second month of the study after the corpus tool had been introduced, but before its use was required (Zodiac Story, Lesson 15, Lesson 16, Exam 15&16). The third section considers the learner logs completed during the second month of the study, once students were required to use the tool in completing in class assignments and the exam (Lesson 16, Lesson 17, Review, Exam 17&18). The final section analyzes the responses given at the end of the study on the end of semester questionnaire.

4.2.1. Month 1: Learner Logs

4.2.1.1. Assignment Learner Logs

During the first month of the study the course continued unaltered, except that students assignments, exams, and learner logs were collected for analysis. As mentioned, students had completed learner logs in the prior semester as well as during the semester of the study before the data was collected. Accordingly, completing the learner logs was something the students were already familiar with. The reflections and analyses present in the logs from the first month of the study are similar to what was observed in log entries previously. Not surprisingly, the main resources students relied on were materials distributed and discussed in class, as well as resources that had been presented in class.
There were a few additional online resources, such as online dictionaries, that are also mentioned that were not presented to the students by the instructor.

Overwhelmingly, students indicated that the primary resources used to complete assignments during month one were the vocabulary list distributed by the instructor, the textbook, and an online dictionary. The vocabulary list referenced by the students was a list that included the characters of the new vocabulary the students were expected to learn during the unit, as well as charts demonstrating the stroke order for writing the character. In class, the instructor would present the new vocabulary to the students and discuss stroke order, pinyin, meaning, and use. Students completed the list in class by adding the pinyin, English meaning, and examples of use. The textbook functioned as a topic guide in presenting each unit. Thus, while much of the material presented in class extended beyond the purview of the textbook, the book was still a valuable resource to view how vocabulary was used, and to review new grammar structures. On the learner logs most participants indicated only generally that an online dictionary was used. A few individuals did list resources such as Google translate, nciku.com, and the mandarintools.com dictionary. In addition to the three categories of resources mentioned, several individuals completed the assignments with their peers and also listed their peers as helpful resources in completing the tasks. Peer work was encouraged by the instructor in completing the assignments, as tasks required students to develop example sentences of their own using new materials discussed in class, and peer mediation was viewed as helpful in assisting students to master new vocabulary, grammar, and other language concepts.
In evaluating the resources used, participants generally concluded that the vocabulary list, notes, and book were the most helpful resources for them in completing assignment tasks. Only one participant, Isabel, indicated that online resources were more helpful. This participant indicated that online resources were more efficient in completing the tasks as online “you can pinpoint exactly what you need”. She went on to conclude that for her, it is usually faster and easier to find the information online that she is looking for. Isabel’s assessment was an anomaly, though, as all other students relied primarily on their notes and textbook. In discussing why these resources were viewed as most beneficial, participants indicated familiarity with these materials. Ellen sums up the general assessment well as she states that in her notes, “I know where to find most of the information I need”.

The general reliance on notes and textbook is not surprising, as participants were familiar with the materials and had grown accustomed to using them in the months prior to the study. When completing assignments, many participants struggled with providing the correct pinyin for characters in assignment 13, and in remembering the meaning of several characters in Lesson 14. For Lesson 13, all responses to what was most challenging about the task indicated pinyin. Most responses generally described the challenge as being identifying the correct pinyin for the characters. However, two responses were particularly insightful in going beyond this general observation to state that the challenge was in remembering the correct tone mark for the pinyin. This reflection is consistent with instructor observation, as on exams students often produced the correct pinyin, but were unsure of tone. The tone mark is often misplaced, and/or incorrect, despite the correct spelling of the pinyin word.
In responding to the prompt to describe a resource that would make the task less challenging, most participants indicated that all necessary resources were already available to them, and that simply more time spent studying and reviewing materials was required. However, a few participants did indicate that having a resource that provided them with example sentences using both new vocabulary and grammar would be very beneficial. One participant, Cathy, referred directly to the challenge of pinyin and tone marks already discussed and indicated that a resource that allowed her to click on any character and view the pinyin and tone marks would be helpful, as this is the information she always struggled to remember, and spent a lot of time looking up.

Encouragingly, in assessing what was most beneficial in completing the assignments 13 and 14, participants identified the same aspects they found most challenging. Specifically, writing out characters, reviewing pinyin and tone marks, and constructing sentences using new vocabulary and grammar constructs. Three participants (Diane, Ellen, and Leo) explicitly mentioned the value of using the newly learned material in constructing sentences using characters and concepts learned from previous lessons, indicating that the new material was internalized and produced in context of previously mastered material.

4.2.1.2. Exam Learner Logs

The exams given to students, both prior to and during the study, focused primarily on production and use of new materials and concepts. Thus, while the first exam task required students to fill in blanks for missing characters, pinyin, and/or meaning, the primary focus of each exam was using the new materials to produce sentences, descriptions, and other unique written materials. In addition to the written production,
each exam also included a dictation section. As the primary focus was on using new concepts to construct meaning, participants were allowed to use online resources and some notes and materials. Participants could reference their vocabulary lists, the textbook, and a few specific online resources including nciku.com, mandarintools.com, and for the exams during months two and three of the study, parallelcorpus.com. The exams were designed to challenge the students, so that even with access to these materials, a certain level of mastery of the concepts was required in order to produce appropriate responses and to be able to complete the exam tasks during the class period in which the exam was given.

The learner log assessments from exam 13 and 14 given during the first month of the study, which were consistent with instructor observations, indicate that the only resources used during the exam were notes and the online resource nciku.com. As with the assignment learner logs, participants used their notes primarily to identify correct pinyin and tone for characters being used. The use of nciku.com during the exam stands out in contrast to the resources used for assignments. Nciku.com has a tool that enables users to draw an unknown character using the mouse as a pencil in order to then search for characters. This appears to be the primary use of the tool in completing the exam, suggesting that the tool was used to identify characters on the exam that students did not yet know or were unfamiliar with.

In self-reporting what was most challenging about the exam, two themes emerged. Three participants indicated that writing the actual characters was most challenging, as several of the new characters had “small strokes that are difficult” (Cathy). The general consensus, though, was that constructing grammatically correct
sentences was the most challenging. Reasons for this included, not knowing many adjectives, remembering the proper use of verbs such as 是 (to be), and feeling unsure about word order. Suggested resources for making the task less challenging all included hypothetical resources that would address these exact issues.

4.2.2. Month 2: Learner Logs

4.2.2.1. Assignment Learner Logs

The parallel corpus tool was introduced at the beginning of the second month of the study. The first assignment completed during this portion of the study, the Zodiac Story assignment, was adapted to also provide participants with experience in using the parallel corpus tool. For this assignment students read through an abbreviated version of the story of the 12 zodiac animals using the parallel corpus and then completed a series of lexical, grammatical, and composition tasks. Students also completed Lesson 15 and Lesson 16 assignments during this portion of the study and had the option to use the parallel corpus tool, but were not required to do so.

For the first experience with the parallel corpus in completing the zodiac story assignment, participants reported overall positive, but mixed reactions. The primary resource used to complete the assignment was the parallel corpus tool, but participants also reported using their vocabulary lists and notes, as well as additional web-based Chinese-English dictionaries (though no specific site was named). The most challenging aspect of the assignment reported by the participants was the extent of new or unfamiliar lexical items that were encountered. This is one aspect that participants reported mixed reactions to, as they could not simply look at a character and see what the English meaning was. Instead, they could view the pinyin and then had to work their way
through the meaning of each sentence. Diane reported that she found it difficult to derive “which characters corresponded with the English word, because I had to first figure out the meaning of the entire sentence”. This sentiment is representative of the overall perspective of participants with five participants using the word meaning and an additional four participants using the word understanding to illustrate that they had to ascertain the meaning of the entire sentence before disambiguating the individual words and characters.

Not surprisingly, due to how the assignment was designed, the parallel corpus website was reported as being most useful in completing the task, and participants went into further detail in describing what aspects of the tool they found most useful. The most user-friendly aspect of the tool described was the access to pinyin and tone marks. Three participants provided similar utterances to Cathy who stated that accessing pinyin was very helpful because “sometimes I remember the pinyin and not the character and it was easy to just click on it”. This perception was positively repeated as participants recognized they could work their way through a story without needing to spend excess time looking up characters. Leo was the only individual to extrapolate beyond this aspect of the tool, and he described the search function being most useful to him as he was able to “look up words that were confusing and [he] could then understand the sentences”. Beyond using the tool just to read and access the pinyin, he also used it to search unknown terms and then viewed the resulting concordances to derive meaning.

The only issue found lacking in the tool based on the learner log responses was the ability to “distinguish which characters formed a word”. Ironically, Isabel listed this aspect of the tool as being the most beneficial resource. The tool does, actually, highlight
which characters combine to form a word as the user scrolls over the characters. This
difference in student perception and experience is understandable given that this was the
first time participants had extensive experience with the tool, and were still learning its
function and features.

The responses describing the corpus tool in Lesson 15 and 16 learner logs are
consistent with the observations made from the zodiac assignment, but provide further
detail as the participants experience and knowledge of the corpus tool progressed. In
contrast, though, participants did not rely primarily on the corpus tool in completing these
two assignments, but rather describe it as an additional resource to the vocab list, notes,
textbook, and other online resources they were already familiar with. The general
responses on these learner logs were consistent with what was observed in the Lesson 13
and 14 learner logs. Participants found their vocabulary lists most helpful in addressing
the tasks of the assignment, but referred to the textbook and online resources when
needing further assistance.

Also similar were the primary challenges participants reported struggling with
including correct pinyin for characters and struggling with innovative sentence
composition in using new vocabulary items and grammatical structures. Three
participants who did not report using the parallel corpus to complete these assignments
(Cathy, Diane, and Gail) described their struggles with sentence composition much the
same as they did during the first month of the study. Gail reported using only the
vocabulary list in completing the assignment as it helped her with identifying characters
and writing the correct stroke order, but she reported struggling with “having the correct
order of words so that it fit the correct sentence structure”. Ellen, Hanna, and Isabel, on
the other hand, each reported using the parallel corpus to complete the task. While they each reported similar struggles in creating sentences, they also reported that the parallel corpus was a great resource in overcoming this obstacle. Hanna stated, “nciku helped me find what new vocab items meant, but parallel corpus helped me figure out how to use them”. Ellen and Isabel reported similar experiences in reporting how the corpus tool helped them in writing new sentences and feeling confident about sentence structure and word choice.

4.2.2.2. Exam Learner Logs

While the content for the second exam was different, the format and structure were very similar to the exam given during the first segment of the study. The first section required participants to fill in either the missing character, pinyin, and/or meaning for a series of vocabulary items, the second task was a dictation task, and the remaining tasks prompted unique sentence, and writing composition. Participants were allowed access to the same materials as the previous exam, but with the addition of the newly introduced parallel corpus tool. Participants’ reported experience mimics that which was reported in completing Lessons 15 and 16, with many of the students experiencing the same challenges and tasks as previously reported, but without any changes in approach. Four students, however, reported using the parallel corpus tool to address their challenges.

As with the first exam, students reported using their notes and vocabulary lists to address problems with identifying the correct pinyin and tone for characters they encountered. Similarly, nciku was reported as the primary source used for identifying characters students were unfamiliar with, again using the drawing feature of this tool to
draw the unknown character and then identify its pinyin and meaning. In contrast to the first exam, four students reported using the parallel corpus to both search for new vocabulary items, as well as to derive how to correctly use those items. Additionally, two of the four students that used the tool also reported using it to check the word order of their sentences. Leo reported that he used the tool “because [he] could see sentences with the words in them when [he] looked them up”. Diane also noted that this was helpful, and that she appreciated being able to see several example sentences and “the translation, pinyin, and character for all the words [she] needed”. Ellen added that it was difficult “coming up with enough sentences for the last part, but still having variety,” but that “parallel corpus helped a lot with this”. The self-reported observations of these three and Gail, suggest that at least four of the participants were using the tool as instructed to analyze authentic language samples to then derive both meaning and use of the lexical and grammatical items they were struggling with.

4.2.3. Month 3: Learner Logs

4.2.3.1. Assignment Learner Logs

During the final segment of the study, participants were required to use the parallel corpus tool to complete several in-class learning activities, as well as in completing the exam for this unit. They were also strongly encouraged to use the parallel corpus tool to complete the three assignments given during this section. All of these assignments were begun in class and then finished as homework. All participants were observed using the parallel corpus while working on the assignments in class, and all who submitted learner logs also reported using the tool on these three assignments. Additionally, the traditional resources cited in the previous sections: notes and
vocabulary list, textbook, and additional online resources were also listed in this final section.

A few of the participants reported that their notes and vocabulary list were the most helpful resources in completing the these tasks, and these same students had little to say about the parallel corpus tool, and the remainder of their learner logs provided little or no detail or description. The remaining five participants that completed learner logs for these assignments all listed the parallel corpus tool as most useful in completing the task. Independently, this observation is not remarkable, as the emphasis made by the instructor on using the corpus tool would anticipate an outcome where most students implemented it in completing the task. What is worth noting is the more descriptive feedback describing what in particular was useful and/or lacking in using the corpus tool.

As with prior assignments, participants highlighted sentence composition as being a challenge for the assignments completed during the final segment of the study. This particular challenge was even more evident during this section, as students were required to compose small essays and not simply individual sentences. All eight individuals who responded indicated that grammar and sentence structure were the biggest struggles. While three found their notes and vocabulary list adequate, the other five provide insight into how they used to the corpus tool to navigate this issue. The general observation was that the tool made it easier to decipher the Chinese by easily recognizing pinyin and then identifying the characters. Four of the participants mentioned that being able to search using pinyin and/or English made it much easier for them to find and remember the individual characters and words they wanted to use in composing their writings. Additionally, three of the participants also described being able to search for a term and
then view it in several example sentences as being very beneficial in learning to use the term grammatically correct and in understanding its meaning. Kate echoed these sentiments but did provide some criticism, stating that it would be more efficient if one could “just search English, pinyin, or Chinese instead of separately” when using the corpus tool. As the parallel corpus currently exists, the user must select which language format is being entered every time a search term is entered. Kate found this frustrating and less efficient than other online resources.

4.2.3.2. Exam Learner Logs

The Lesson 17 and 18 exam followed the same structure as the previous exams, but the final section required participants to complete a small essay at least forty characters in length. This was the longest writing task participants had been required to complete on an exam, and unsurprisingly this task was also what was reported as being most challenging. The learner logs indicate that nciku.com, vocabulary lists, and parallelcorpus.com were the three resources used in completing the exam. Participants indicated that the first two resources were used primarily to look up pinyin for vocabulary items and nciku to look up characters written on the exam that they did not know or remember. The parallel corpus was reportedly used primarily to derive sentence structure.

Leo was particularly descriptive in his learner log and indicated that he struggled with “the story because it was challenging to come up with what to write about, and two [exam] questions because they had strange sentence structure and I forgot some of the words”. He then described using the parallel corpus to look up words from the exam questions to figure out their meaning so that he could respond appropriately. He
concluded stating that with the parallel corpus he had “everything that [he] needed” to work through the exam. Gail came to the same conclusion, but based on her use of the tool to identify and “double-check that I used the correct pinyin and tone marks”. She indicated that she found the tool to be the most proficient resource in looking up tone marks. These responses were particularly positive, but all participants reported having adequate resources to navigate the exam. There was no criticism offered of either the corpus tool or other materials, with participants concluding that they “[couldn’t] think of anything else [they] would need,” and “the sources are enough”. As noted, these final learner log assessments were without negative comment, and provided no new suggestions for additional resources for future tasks.

4.2.4. End of Semester Questionnaire

The end of semester questionnaire was a more detailed reflection of the participants’ general experience in the class, and specifically with the parallel corpus. The questionnaire asked participants to respond to eight open-ended questions that encouraged narrative responses evaluating their overall experience in the class. The second half of the questionnaire pressed participants to analyze their experience with the parallel corpus. The first three questions asked: 1) how useful the parallel corpus was, 2) how often individuals used it, and 3) how likely they were to continue using it. These three questions were answered using a Likert scale, which ranged from one to seven. The remaining five questions were additional open-ended questions that allowed participants to go into detail in evaluating the parallel corpus tool.

The following analysis of responses is divided into two sections in correspondence with the questionnaire. The first section considers participants’
evaluation of their overall experience in the class, but also draws attention to comments referencing the parallel corpus. The second section then explores the explicit participant feedback given about the parallel corpus, highlighting both the positive aspects that individuals identified, as well as the negative or lacking features that participants feel are still wanting in the corpus tool.

4.2.4.1. Course Evaluation

The first questions prompted participants to consider what aspect of the class they found most enjoyable. The most common responses described the in-class learning activities, and the small class size. Gail described the class as “a really fun learning environment where everyone interacts”. Cathy made a similar observation, “The class size, I was able to learn a lot more because there weren’t as many other kids trying to learn as well”. This response is intriguing as the participant recognized other students as being potential roadblocks to her learning, but felt that was not the case in the Chinese class. Overall eight of the 12 participants described the interactive classroom learning activities as being the most enjoyable, and five attributed the success of these activities in part to the small class size. Overlapping with this topic, four participants highlighted speaking and using the language as being what they enjoyed most, as “speaking Chinese is fun and I started to comprehend the language” (Kate). Three participants (Diane, Ellen, and Leo) all identified reading and writing characters as being their favorite aspects of the course. Leo described lessons “learning about the characters” referring to lessons about character structure, stroke order, and etymology, while participants Diane and Ellen both enjoyed learning to use the characters, and “being able to read and write the characters that we learned” (Diane). While these few participants identified writing
as the most enjoyable aspect of the course, the overwhelming majority then responded to
the next prompt by recognizing this topic as being the most difficult.

Ten of the 12 participants felt that learning characters and reading and writing
Chinese was the most challenging aspect of the class. The general consensus was that
writing characters is complicated and involves a lot of memorization. Additionally,
seven participants also explicitly mentioned memorizing the tones as well as the
composition for each character. As Hanna described it, “I found it hard to memorize the
characters and tones. There wasn’t a way to learn them besides just practicing and
memorizing”. Besides this struggle, individuals also mentioned grammar structure (2
participants), pinyin and pronunciation (3 participants), and composing sentences. The
two individuals that did not mention reading and writing (participants Brian and Frank)
both identified time as being what challenged them most. Specifically, they found it
challenging that the class did not meet daily, and they did not have daily contact with the
language.

As Norman (1988) described, characters are challenging, but also intriguing to the
foreign language learner. Thus, while the majority of participants explicitly highlighted
their struggle with character acquisition, several of them also identified lessons on
characters and reading comprehension as being their favorite lessons. Diane stated that
she “loved reading stories in Chinese,” and identified the parallel corpus tool as making it
possible for her to “read stories, and actually understand them”. Cathy and Leo shared
this sentiment, and both described their intrigue and love of characters. Five participants
identified lessons on or about characters, reading, and/or writing as being the most
enjoyable. Two participants identified one specific classroom session that was on colors,
and the remaining five participants all described the lessons on food (Lesson 16). This topic proved to be popular as participants were able to experience and not just learn about the topic, as there was a class field trip to a local Chinese eatery, and a guest who later came and taught the class to cook a traditional Chinese dish.

As participants evaluated the resources they found most helpful in studying the language and then described how they used them, the same four primary resources mentioned in the learner logs all resurfaced: vocabulary lists and notes, the textbook, the nciku.com website, and the parallel corpus tool. The corpus tool was identified most frequently, with eight participants referencing it. While this observation is not independently remarkable, as students were required to use the tool for part of the class, the detailed feedback provided is worth exploring. Gail, Kate, and Leo provided nearly identical responses as they described why the parallel corpus was useful. In Leo’s words, “you can find the answers you need, but you still have to search for it and figure out what it means”. Gail went on to provide more detail on how she used the tool, “I liked parallel corpus because you could read through the English to get an idea of what something meant, and then work through the Chinese. I also liked how you could select each individual character and work through a sentence”. In addition to using the parallel corpus as a search tool to identify specific terms and their uses, other participants also mentioned the tools benefit in learning tones and pinyin. Leo stated that he also “always used it for tones. That repeated action of looking and using tones put them in my head”. Several participants mentioned that they liked having a resource that had the characters, pinyin, and English all in one place.
The nciku website was identified by five participants, and was also viewed as a very valuable resource. All participants who mentioned this resource went on to describe the drawing feature used to search for unknown characters. Brian stated, “I used nciku to draw characters and get the pinyin translations”. While participants were observed using several other of the useful features of nciku (see section 4.3.2.), none of these other features were identified on the questionnaire. Five participants also identified their vocabulary lists and notes as being very helpful. Cathy stated that she relied on the vocab list, “so I knew I was drawing the characters correctly”. Diane, Ellen, and Gail made similar comments but referred specifically to referencing the list to learn stroke order. Diane also stated that she found her notes the most useful resource for correcting and learning sentence structure. The final resource mentioned, the textbook, was only mentioned by Leo who stated he liked using it together with parallel corpus “for different activities”.

The final prompt on this section of the questionnaire asked participants to suggest changes to improve the class. Seven participants desired more class time, and suggested that the class meet every day, as opposed to meeting only twice a week under the current structure. Two participants desired more listening and speaking practice, and Jack felt that there should be fewer vocabulary terms, focusing on a couple characters per lesson instead of the typical 6-12. Lastly, Cathy, who consistently referenced her use of the vocabulary list, suggested that example sentences for each vocabulary item be added to this resource.
4.2.4.2. Parallel Corpus Evaluation

The latter half of the end of semester questionnaire first prompted participants to analyze the usefulness of the parallel corpus tool by rating it on a scale of one to seven for a) how useful they found the tool through the semester, b) how frequently they used the tool, and c) how likely they were to use it in the future. The lowest response given was a four out of seven, and the mean and mode for each question can be viewed in Table 4.3. Due to the limited number of participants (12), these ratings do not serve to independently demonstrate the usefulness of the tool, but rather suggest the perceived value of the tool among participants. With responses to all three prompts ranging between five and six, the ratings suggest that participants valued the features of the parallel corpus tool for learning Chinese, and plan to continuing to use the tool in the future.

Table 4.3.

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean</th>
<th>Mode</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - How useful was the parallel corpus tool</td>
<td>5.67</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>for you throughout the semester?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 – How often did you use the parallel corpus</td>
<td>5.08</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>tool in learning and studying Chinese?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 – How likely are you to use the parallel</td>
<td>5.67</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>corpus tool in the future?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The final questions, then gave the participants an opportunity to analyze the tool, and provide both positive and critical feedback. Question four asked participants to describe what they liked most about the tool. The primary response given referred to the different language formats being accessible all in one place. As Cathy stated, “I found the side-by-side English, pinyin, and characters VERY helpful. I could go through and
see how vocab was used in context”. Gail highlighted the fact that one could see these different formats all in one spot without having to click to different screens, and Isabel found this very helpful for reading through texts and being able to quickly identify characters she hadn’t learned yet. Ellen’s response also indicated how she used the tool in practicing and acquiring characters. She stated, “I liked how you could click on a character and it would tell you the pinyin but not the English. That way, if you knew the pinyin you could think through what the sentence meant”. Her response indicates that she used the tool as it was intended – to focus on character acquisition and to use the additional language formats as resources, not as crutches. Collectively, responses indicate that having the multiple language formats all in one place was the feature that participants appreciated most in using the parallel corpus. Additional responses specified that the simplicity and “easy usability” of the tool was also appealing (Jack).

These responses were reiterated in the answers given to question 6 that asked individuals to describe how they used the parallel corpus. Most responses again mentioned using the language formats, but then went further in describing how this feature was used to learn new vocabulary, grammar structure, and composition. Ellen plainly explained that she used the corpus “when [she] was looking up specific characters and when [she] was seeing how to form sentences”. Hanna explained that she mainly used the pinyin when she had trouble reading a character, but would also refer to the English when she couldn’t figure out the meaning. Frank’s response was unanticipated as he mentioned using the tool to “see how Chinese is written in big paragraphs”. He continued on to explain that other course materials exposed him to words and sentences, but the parallel corpus allowed him to see the Chinese language written in paragraphs and
essays. Several participants also indicated the value of the tool in learning and reviewing tones. Leo stated that he liked the tool “for the tones when [he] clicked on the characters”. He further clarified that he liked being able to view just the tone to verify pronunciation, without the “pinyin being given to [him]”. The majority of these responses indicate that participants were using the tool, or at least describing their use of the tool, as a mediational tool and language learning resource in accordance with how they were instructed to use it when it was first introduced.

Criticisms of the tool in response to question 5 identified both limitations of the resource as well as some glitches. Cathy described one primary issue, which was also mentioned by three others. She criticized “how scrunched the text got when both the English, pinyin and characters were on the screen”. This was an issue that became apparent only after using the tool in class. When only two language formats (characters and pinyin OR characters and English) were visible on the screen, the content was clearly aligned and easy to follow. However, when characters, pinyin and English were all simultaneously opened, there was a formatting issue on some devices that misaligned the text and made it hard to follow. Besides this shortcoming, two participants also found fault in not being able to see exact correspondence between characters and English words when reading a text. When a specific word is searched in the tool, word alignment is highlighted, however, when viewing sentences or complete texts there is not a feature that allows the user to identify word-to-word correspondence between the character and the English. This design feature was intentional so that users would be forced to analyze sentences to derive meaning and not be provided with word for word translations, but it was evidently a feature that some participants disliked. The third critique of the tool was
the lack of traditional dictionary features. Consistent with the previous criticism, two additional participants disliked having to derive meaning from sentential context, and would prefer if an interactive feature providing definitions were added. The final criticism offered identified the current limited size. The two individuals offering this opinion indicated the tool was quite useful for the course, but for it to be of value to them in their future study of Chinese more material would need to be added.

Question 7 then offered participants an opportunity to respond to their own critiques by providing suggestions of how to improve the tool. The same two individuals that critiqued the lacking dictionary feature, also then suggested that this feature be added. Similarly, adding a wider variety of content was also recommended. All other responses, though, proposed adding multimodal features to the corpus tool. Several individuals would like to see audio features added to the tool. Specifically, it was mentioned that in addition to the interactive feature that allows the user to view tone mark above the character and then pinyin, that an audio pronunciation would be very useful. Furthermore, Leo recommended that a complete audio recording for all stories and essays be uploaded. Hanna provided the final recommendation, that pictures and short video clips also be added to both aid in learning content as well as to make the tool more interesting and interactive. Overall, the critiques and suggestions provided are very constructive and many of the recommended features would undoubtedly improve the parallel corpus tool, and likely the user’s learning experience. However, the evaluations of the tool also indicate that participants value the current features of the tool, and were able to identify specifically how the parallel corpus aided them in their acquisition of Chinese.
4.3. Instructor Autoethnography Analysis

The introduction and adaptation of the parallel corpus tool into the classroom was as much of a learning experience for me, as the instructor, as it was for the participants. Being a non-native speaker of Chinese, I anticipated how the tool might be used and adapted by the participants, based on my own experience in studying Chinese, but there were many unknowns going into the study. In order to track my learning progress as an instructor, and to document how the technology was adapted into the classroom, an autoethnography was maintained during the study. One journal entry was made for every class period during the duration of the study (22 in total), by responding to a set of ten questions. The purpose of the autoethnography was to a) document how the instructor prepared for each class and prepared to use the parallel corpus tool in class, and b) to provide a post-instructional account that detailed how the class, and implementation of the corpus tool, went. The ten questions prompt responses as to whether the tool was used, how it was accessed, how the instructor prepared to use it, adaptations that were made, learner experience, challenges, and other thoughts and impressions. The detailed prompt used to complete the autoethnography responses can be viewed in the Appendix. The following is an analysis of the instructor autoethnography and, as with the learner log analysis, is divided into three sections in accordance with the three segments of the study.

4.3.1. Month 1

The parallel corpus tool was not adapted into the classroom during the first month of the study, and the responses to the prompts of the autoethnography during this portion of the study are limited. Question 4, which asks “how did you prepare to use the tool for the class” and question 7, which asks “what was the learner experience like” were the
only questions addressed during this portion of the study. The responses given during this time primarily focused on how the instructor was preparing to use the tool in the coming lessons, and reflections on how the tool might have been beneficial in covering the current days lessons. There were six class periods during month 1 of the study, and two of those class periods ended up being snow days, where school was canceled due to adverse weather.

Preparations that were made during this time primarily consisted of preparing materials that would be used during months two and three of the study and making sure that the material was correctly uploaded to the website. The work that was spent preparing the material was a necessary aspect to both prepare for the class periods in which the parallel corpus would be used, and to make sure technological ‘glitches’ were worked out of the system. On average, it was reported that 1-2 hours were spent in preparing material each time a journal entry was made, with longer work times reported on the two snow days when class was canceled. The process taken to prepare and upload materials to the parallel corpus tool was somewhat labor intensive, and involved several steps. The material on the website is segmented at the word level, sentence level, and separated into individual documents. Accordingly, all text that was uploaded to the parallel corpus had to be prepared in three different formats and then uploaded separately.

To prepare the individual documents for upload, the instructor would begin with a .txt file in Chinese and then using Unix programming, run the file through the Stanford Chinese word segmenter. Word segmentation was needed so that the corpus tool could later identify and highlight individual words for the user. After completing this step, the instructor would then convert the file to pinyin, and then upload the Chinese character,
pinyin, and English files to the parallel corpus interface. The second step required separating the documents into sentences. Again, using Unix programming, the sentences were segmented, and then Chinese character, pinyin, and English sentences were aligned in a table and then uploaded to the website. For the final step, Unix programing was used to create word lists from the files, and then align the word lists into the same three formats: Chinese character, pinyin, and English, as well the additional Chinese character plus tone mark format. This final format was achieved, again through Unix programming, by indicating the tone mark by number instead of diacritic (see Table 3.2), and listing the character followed by the tone mark number, delineated by a comma. The four formats were then aligned in a table and uploaded to the website. As most work was completed through programming, the size of the document did not affect how long the preparation process took. What did take longer, though, was proofreading through the final alignment to make ensure the final product was correctly aligned and functioning. On average, the instructor took one to one and half hours to prepare and upload each file to the corpus. This process was somewhat labor intensive, but the majority of the work was completed before the beginning of the implementation of the tool and did not conflict with the instructors preparations for classroom learning activities.

In response to question 7, the instructor did anticipate the value of the tool for the students as they learned and practiced new vocabulary and grammatical items for the current lessons they were working on. Specifically, two accounts are given in which the instructor detailed how the corpus tool could have aided the day’s lesson. For Lesson 13 participants struggled with tones in pronouncing ages. Participants had previously learned numbers and were confident in expressing numbers, but rarely managed to utter
the correct tone. In the journal entry for the day it was noted that, “a small reading activity where they could click on the character and view the tone every time they said a number would be very helpful, and make it possible for students to practice in pairs and self-correct. This would allow them to get more individual practice instead of having to practice as an entire class, so that I draw attention to the correct tone each time a number is stated”. While this activity could be completed using other methods, the comments reveal how the instructor was planning to use the tool, and looking for opportunities to incorporate it to aid the participants learning experience.

The second observation noted by the instructor in association with Lesson 14 actually evolved into the first activity the participants completed using the parallel corpus tool. The topic of Lesson 14 was animals, and students learned the names and characters for several different animals. The instructor noted that while the students were retaining how to say the names of the various species, they had trouble reading and remembering the characters. While some characters such as 猫 (cat) and 猪 (pig) have the same radical and have some resemblance, participants were also mistaking characters such as 狗 (dog) and 马 (horse), which are markedly different. In his journal, the instructor noted that a reading exercise that required participants to understand and use several animal characters would be helpful, so that participants could actually use the terms and focus on meaning and use, instead of just memorizing a vocabulary list. The challenge for a reading assignment, was that participants overall vocabulary was still limited, but the instructor noted, “this seems like the perfect opportunity to introduce the parallel corpus, so that they can read through a text that would otherwise be too difficult for them”. This
final entry is the last note made during the first month of the study, and also the first activity of month 2.

4.3.2. Month 2

By the second month of the study participants were aware that a new online resource would be introduced and used in class, but they had not yet had the opportunity to view or use it. Based on the experience and reflection at the end of month 1, the instructor planned a lesson on the 12 zodiac animals to be the first lesson introducing the parallel corpus tool. An abridged version of the story was uploaded to the online corpus, and for the class the participants were to read through the story and then complete an assignment that checked for reading comprehension and then required them to use the animal vocabulary to produce sentences. While the parallel corpus was basically prepared and ready to use, the addition of this new activity required the instructor to devote a couple of additional hours in preparing the text and uploading it to the website.

All preparations were made, but the first time using the tool ended up being full of challenges. For the first time during the school year, the school’s server crashed right before the class began, and the server remained down for the rest of the day. One of the objectives of the activity was to familiarize the participants with the online corpus, and as such the activity and assignment required online access to the parallel corpus tool in order to complete the project. As the server was down, the lesson plan had to be abandoned for the day, and an alternative activity developed instead. Instead of completing the exam, the class went outside and focused on speaking activities to use the vocabulary they have to describe their surroundings and play a few visual description games. While the class ended up being an overall success where participants were able to review and use their
vocabulary, the issues with technology required a drastic change in plans. After this experience, the instructor recorded in his journal entry, “The challenge of technology – when it works, it’s great – when it doesn’t you’re left stranded”. He further stated, “I think it will be necessary to always have a backup plan in place when using this or any other technology – whether this be a single default lesson or a way of delivering the same lesson without technology”. While the first day of using and introducing the parallel corpus tool did not go as planned, it certainly served as a reminder of the necessity for preparing alternative plans when using technology in case of events such as this one.

While technological tools such as the parallel corpus can function to mediate learning an advance development, if alternatives are not prepared for if the technology does not work then learner development is hindered instead of helped.

Fortunately, by the next class period the server issues had been resolved, and the Internet – and access to the parallel corpus tool – were fully functional. The instructor did note in his journal that in preparation for the lesson he had made several copies of the zodiac story, so that the lesson could proceed in case issues with the server or technology still persisted. Other than making the handouts for an alternative to completing the activity, no additional preparations were made or required. The instructor noted that the activity went as planned, and ended up being a positive learning experience for all. As this was the first experience with the corpus tool, the instructor had each participant record their screen to be analyzed later. No additional challenges were indicated, but it was noted that the participants seemed to enjoy using the tool. He recorded, “the students seemed to enjoy the activity and were quick to point out that there were many words and characters that were new to them – but that they were able to figure out meanings all the
same”. The concluding remarks indicated that the lesson plan met and exceeded its objectives in assisting students in acquiring characters and new vocabulary on the topic of animals, as well as to familiarize them with the corpus tool and help them begin to feel comfortable with it.

For the remaining two class periods before the second exam, lesson plans were consolidated to make up for instruction days lost due to snow days and the server crashing. Lessons 15 and 16 were covered in just two class periods, instead of four. Accordingly, the instructor noted that the class preparations focused more on discussing and practicing the new material, and participants were given less time than normal to work on their assignments and practice independently in class. Consequently, this also resulted in less time spent using the tool than had initially been anticipated, but the participants did still have opportunities to explore and use the tool in class and for their homework. The one note made by the instructor was that several of the new grammatical constructs were still confusing to students when using the parallel corpus, and additional mediation from the instructor was required to help students comprehend what was being expressed in the content on the parallel corpus. The instructor noted that this additional assistance seemed to be equal parts instruction on the grammar and content being viewed, and on how to use the parallel corpus as a tool to decipher the meaning and content. In other words, the instructor made sure the students understood the content of the lessons, but also instructed them on how to better use the tool, and evaluate the language themselves.

The final entry from this portion of the study concluded with exam 15/16. For this exam the participants had access to the parallel corpus, as well as other online
resources. The instructor noted that his biggest concern in preparation for the class period was encountering another glitch with technology. As backup, he planned to postpone the exam, and take an additional day to review and practice the material. No back-up was needed, however, as the class proceeded as planned, and participants computer screens were again recorded as they completed the second exam. The instructor noted that while all students did use the parallel corpus at least once in completing the exam, many individuals still used the resources they were already familiar with. In particular he noted that the website nciku.com was used to look up words, phrases, and pinyin, that he had anticipated the participants would want to use the parallel corpus for as the material on the corpus provided more explicit samples and only materials with which the participants were familiar. He concluded that while some participants were beginning to see the value of the tool, more instruction on how to best use it was needed in order for the parallel corpus to function as a valuable mediational tool in their Chinese language development.

4.3.3. Month 3

For the third month of the study, participants were required to use the parallel corpus tool to complete the assignments. Two class periods were committed to each lesson, and the first day of each lesson focused more on instruction of new materials and the second day focused on practice and in-class assignments. For both of the instructional days there is little recorded about the use of the parallel corpus in the autoethnography, as its use was minimum. During these days the instructor did, however, use the tool to display new vocabulary items and grammar constructs on the projector. As all of the textbook material was also included in the parallel corpus, the
instructor used it as a platform to introduce new materials to students, who could then either follow along in their textbook or online at their individual computer. The reflection of these experiences indicated that the tool worked well, and as the material was already uploaded to the corpus it was easier for the instructor to use the tool for teaching than to prepare an alternative media presentation or write on the board.

During the second day of each lesson when the participants practiced material and began working on assignments, the parallel corpus was used extensively. Lesson 17 was on family and focused on learning kinship terms and discussing family members. As described previously, (see section 3.3.1.4.), the students used the parallel corpus to read a letter in which the author, Xiūmíng, described the members of his family. The letter primarily consisted of new vocabulary items learned in the previous class period, but also had some additional terms and grammar concepts. The assignment required the participants to explain the meaning of the letter and to identify new vocabulary items and describe their meaning and use with only the corpus as a resource. As homework, participants wrote their own letter back to Xiūmíng describing their own family, incorporating the vocabulary terms they had just identified. In preparation for this class period, the instructor noted that he again printed out some paper copies of the letter, so that participants could at least begin the task if technological problems were encountered. He also recorded that in preparation for the class he used the parallel corpus to search the new terms that participants would encounter when reading the text, in order to make sure several example concordances would be returned. The main concern mentioned was that he wanted students to be able to search for terms and derive meaning from the
concordances produced from the search. In his preparations, he found adequate examples, and did no further preparations.

After the class period he reported that though this was his main concern about the activity, it did not end up being a problem for participants. Though they struggled with a few of the concepts and required additional mediation from the instructor, the data returned from the tool was adequate. He reported that the learning experience was very interactive, and that participants worked together in reading the text and using the parallel corpus more effectively. He also noted that two participants (Ellen and Isabel) used the tool to produce their written responses. In addition to using the assigned text, the instructor observed these two participants using the concordances from terms they had searched to create their own sentences. In other words, they modeled the structure of their sentences on the example concordances from the tool when using new terms and constructs. One final comment noted, was that a problem with the program was identified. The pinyin search feature only worked from the home screen. Therefore, after completing one search participants could not then enter a second term and search again, but had to return to the home screen and clear their previous search. This issue was addressed by the software programmer and corrected before the next class period.

For the Lesson 18 activity and assignment, which was on occupations, participants first interviewed classmates about their dream job and the occupations of their family members, and then completed an assignment in which they wrote an essay describing their personal dream job, and their family members occupations. In preparation for this assignment the instructor noted that he was most concerned that participants felt confident using the parallel corpus as a resource for their writing, as this
assignment was the most challenging writing assignment they had encountered. Specifically, he planned to provide examples of how to look up new grammar constructs using the tool and evaluate the sentences displayed, so that participants could use the corpus as a resource to mediate their sentence composition. As participants began working and using the tool, he noted that he observed several participants using the tool as just instructed, and that almost every individual also used the tool to search for new Chinese vocabulary terms for occupations that had not been taught in class. Since not all terms searched for were included in the corpus, many participants then looked to other resources such as Google translate and nciku to find specific terms such as marine biologist. The final comments recorded by the instructor noted that besides the limitation of data available in searching for a wide variety of new terms, the activity proceeded as planned, and it appeared as though the “parallel corpus functioned as a mediational tool in students writing and sentence composition”.

The instructor was absent from class on the day of the review activity, but mentioned that in preparing materials for the substitute instructor he prepared a handout of review activities in preparation for the exam with directions to use the parallel corpus, and any other resources needed. The substitute instructor reported the class went well, and student feedback on the following class period, which was the exam, indicated they felt prepared for their last exam. The journal entry on the day of the exam noted that participants “seemed to use the parallel corpus tool more than their other resources,” but that the same additional resources – nciku.com, vocab list and notes – were still used in completing the exam. The instructor also remarked that in contrast to the previous exam, he observed several individuals using the parallel corpus to identify pinyin and tone, and
that many participants search out grammatical structures or terms to “hopefully model their sentences after”.

This exam was the last class period in which the parallel corpus was used in the classroom. There are, however, a few observations from the instructors remaining journal entries that are worth discussing. The next class period was an interactive review of the content from the academic year, and was designed to be an enjoyable last day of review and instruction. At the end of the instructor’s entry, he notes that during the activities three different participants approached him and asked if he would upload a specific story to the parallel corpus so that they could read it over the summer. This indicated not only that the participants planned to continue using the parallel corpus, but also that they felt it was a desirable method for reading Chinese texts. During the next class period participants gave final presentations in Chinese on a topic of their choosing about the Chinese language or culture. Participants were required to have a PowerPoint presentation, and could use notes if they wanted, but all material had to be submitted to the instructor. In reviewing the submitted materials the instructor recorded that he was surprised to find that the three participants who used and then submitted their notes had all prepared them in the same format, which resembled the parallel corpus. The notes were in three columns: Chinese characters, with some tone marks indicated with a diacritic, pinyin, and the English. The only instruction given for the assignment was that it must be all in Chinese, be three to four minutes long, and any materials used had to be submitted. There was no direction given on how to prepare materials, but participants independently prepared their presentation notes following the format of the parallel corpus tool.
CHAPTER 5
DISCUSSION

The analysis of data provided a descriptive overview of how the parallel corpus was incorporated in the Chinese language classroom. The data analysis reveals how participants applied the corpus, how it aided them (or not) in their development, and insight into their individual perspectives of this resource through their learner logs and questionnaires. Furthermore, the instructor’s experience and insight as recorded in the autoethnography provide an additional perspective on the implementation of the parallel corpus. The discussion chapter now juxtaposes these descriptive observations to derive conclusions in response to the three research questions: 1) Learner Achievement – Will the Chinese/English parallel corpus aid students in learning Chinese? 2) Pedagogical Implementation – How do students and the instructor implement the tool in the classroom and their study throughout the semester? And 3) Learner Attitude/Perception – How do students view the application and implementation of the tool? The chapter is organized categorically to individually address each of these three questions.

5.1. Discussion of Learner Achievement

The first and primary emphasis of the research study was on the developmental achievement of learners through the mediation of the parallel corpus. Specifically, the researcher wanted to discover whether the Chinese/English parallel corpus would aid students in learning Chinese, and then observe how its use affected their development. The discussion of this topic is studied through the lens of sociocultural theory, applying SCT to describe learner development in order to understand if and how the parallel corpus successfully functioned as a mediational tool in developing students’ linguistic
and conceptual knowledge of the Chinese language. Prior parallel corpus studies have analyzed the effectiveness of this technology through the lenses of reciprocal learning (Hunston 2002; Laviosa 2002), discovery learning (Bernardini 1996), and data driven learning (Frankenberg-Garcia 2005; Wang 2001), but not yet through the SCT model. Accordingly, the discussion that follows first addresses how development is defined and understood according to SCT, before then examining how the parallel corpus affected learner development.

5.1.1. Understanding Development through Sociocultural Theory

SCT understands development as a dynamic process in which an individual is able to internalize a concept or idea through social interaction, or as Lantolf and Thorne (2006) state, to transform “externals into personally meaningful experience” (p. 153). Thus the current study defines learner development as an individual’s ability to not just learn a term or concept, but to internalize that term or concept and later be able to create his or her own meaning with it. The parallel corpus is recognized as a mediational tool that functioned as a digital form of interaction to assist learners in language and conceptual development. The instructor, of course, provided additional mediation both in instructing students on new content, guiding their adaptation of the parallel corpus, and helping participants to navigate linguistic concepts in a variety of contexts.

A primary goal of the instructor was to present participants with challenging learning assignments that would necessitate mediation in order for them to resolve the task. SCT posits that a cognitive struggle is requisite for development to take place, and that optimal development takes place in the ZPD, which is understood as the distance between what an individual can achieve alone versus with mediation. In other words, the
instructor sought to provide participants with learning activities that required additional mediation (the parallel corpus) for them to achieve their optimal potential. These assignments and activities assumed a top-down approach to learning, following the CBI model, in which a learner was presented with a concept, and then had to navigate that concept with mediation from the instructor, parallel corpus, and additional resources. A primary advantage of this approach to learning is that students not only learn a concept, but also develop the ability to analyze and apply their new conceptual knowledge in multiple settings. Thus in the discussion that follows, evidence of development is identified in those instances where participants are observed first struggling with a new term or concept, before later evincing a conceptual understanding through their ability to appropriately use the new term or concept in a different setting.

As learners were presented with different challenging concepts, their use of the parallel corpus to navigate these tasks was intended to be methodical as guided by the instructor. An ideal usage of the corpus tool was envisaged where users would rely primarily on the Chinese text, and only refer to other formats as needed to aid their development. The instructor guided participants in using the corpus this way in order to direct their focus on Chinese characters, and mediate their development through the Chinese language. As has been observed, not all participants chose to adapt the corpus in this way, and various learning outcomes were observed. The parallel corpus tool currently does not have built in features to ensure that participants use it in this intended way, and ultimately an individual’s approach to using the tool was dependent on learner motivation and attitude. From an SCT perspective, further mediating participants’ use of the tool, and/or designing features that ensured the intended usage may likely have led to
improved development among a wider range of participants. This matter is discussed further in chapter 6, considering how other approaches and/or features might improve future applications.

The final aspect of consideration is participants’ approach to learning. As identified through the data analysis, three main categories of participants emerged based upon both how participants approached learning in the class, as well as how they adapted the parallel corpus. This categorization is informative to understanding learner development (the product) by providing insight into how the students approached learning (the process). This important distinction was identified many years ago by Hosenfeld (1976) as she researched student-learning strategies and how the activity students perform is often quite different than the activity envisioned by the instructor. In considering the observed development (or lack thereof), understanding the process of if and how participants approached learning and incorporated the parallel corpus in their learning process is a primary focus. Accordingly, it is observed that the way in which the participants completed assignments as well how they chose to apply the corpus led to instances where the parallel corpus: 1) seemed to have little or no effect on learner development, 2) proved to be counterproductive, and 3) evidently improved learning experiences and outcomes. These three results are discussed in detail below.

5.1.2. Indifferent Effect on Development

Among those participants in the poor performers category, as well as some of the individuals in the resistant to parallel corpus category, the parallel corpus was observed to have little or no effect on their learning and language development. This was due to two primary factors: 1) students underperforming in the class and not attempting learning
activities (mainly the *poor performers*), and 2) students using the parallel corpus to quickly identify answers, but not explore concepts (chiefly those *resistant to the parallel corpus*). The shortcomings in successfully implementing the parallel corpus among these learners are discussed in detail in section 5.2. Discussion of Pedagogical Implementation, but mentioned in this section to draw the distinction between method and outcome.

Those in the poor performers category were indifferent to their learning experience both before and after the introduction of the parallel corpus. There was an evident lack of effort in the class, on assignments, and during exams, which leads to the apparent conclusion that those in this category were minimally developing language ability and/or conceptual knowledge. Adam, Brian, and Frank made little to no attempt in the class, but on the tasks they were observed completing (primarily exams), they used the parallel corpus tool not to understand the language but essentially to fill in blanks on their exams. This is evidenced by their emphasis on English when using the parallel corpus, and rarely interacting with the Chinese text. Jack did perform better than the other three, particularly in the second month of the study, but his approach to learning showed little distinction from the others. On the tasks that were completed, the obvious objective was to complete answers, but not to understand the material or learn the language. As analysis of these individuals revealed minimal effort at understanding concepts or ideas, it is difficult to derive any concrete conclusions as to how the parallel corpus affected their learning.

Despite the apparent lack of concern for learning material and concepts, the participants in the poor performers category did demonstrate acceptance of the parallel corpus to complete work and tasks. While this was counterproductive to development in
many aspects, as discussed in the following section, one key observation to acknowledge is that some features of the parallel corpus tool forced participants to interact with the text, and actually understand portions of the content in order to use the tool and complete the accompanying learning activities. In other words, the design of parallel corpora requires a user to interact with text and exhibit a minimal level of comprehension in order to use the resource. This is illustrated through Frank’s use of the parallel corpus to complete the zodiac story assignment. Contrary to instruction, he focused on the English text to search out answers for the assignment. The fact that word-for-word equivalents were not identified in the text, though, required him to analyze the Chinese sentences in order to find the appropriate term that corresponded with the English. Though this does not equate to language development, it does illustrate how using the parallel corpus required a further level of analysis and interaction with the text that likely would not have been provided through other learning methods. This implies that even among students such as Frank, Adam, Brian, and Jack, who appear to be uninterested in actively pursuing their language development, the parallel corpus requires a minimal level of conceptual understanding to even use it. At a basic level, using parallel corpora presents learners with complete conceptual knowledge in both their L1 and their target language, which allows them to decipher and navigate the unique features of each language in expressing a particular meaning. At this level, it could be argued that by simply using the parallel corpora some conceptual development is inevitable, as the user is exposed directly to language use and has no choice but to navigate concepts in order to use the resource.

At a second or deeper level parallel corpora also expose learners to translated texts that are in themselves models of how conceptual knowledge has already been
applied (how a translator applied his or her conceptual knowledge of a language in creating the translated text). This inherent makeup of parallel corpora allows students to view texts and concepts in their original form, see how a translator ‘re-contextualized’ the original meaning or concept in the target language, and then further ‘re-contextualize’ the meaning being expressed in forming their own understanding of how meaning is expressed in both their L1 and their target language. Negueruela and Lantolf (2006) identify this “re-contextualization of concepts” as an essential feature of language development that allows students to adapt their knowledge in future settings. Thus from simply interacting with the resource Frank observed how a translator re-contextualized the content, although there is not further evidence that he, or others in the poor performers category, went the next step to re-contextualize the concept in producing their own meaning.

A further observation of how the parallel corpus could potentially improve learner achievement is through the pinyin search feature. Brian and Adam were both observed using the pinyin search feature of the parallel corpus to more efficiently identify unfamiliar characters. Their use of this feature indicated that they recognized this feature of the tool as being more efficient than other resources in searching for a term when they knew the pinyin and tone mark. Though this does not indicate development, it does demonstrate more efficient interaction that could lead to development, as it requires both the mastery of tone marks as well as placing emphasis on identifying the character without spending unnecessary time searching through other resources. This use of the corpus tool is highlighted here, as Diane also used this search feature in the same way, but with a different result (section 4.1.3.2.). Diane completed a pinyin search the same as
the others, actually searching the identical term as Adam, but then went on to analyze the sentences in the search results to then create her own meaning using this term. This next step demonstrates how this and other features of the parallel corpus successfully mediate development and understanding, but only when the user takes the appropriate route in pursuing that conceptual knowledge.

Initially, Hanna, Isabel, and Kate from the *resistant to parallel corpus* category demonstrated similar indifferent results in using the parallel corpus. A key difference is that these participants completed their work and performed well in the class. This good performance is not the same as conceptual development, though, as there was little indication of mastery of concepts – rather mastery of completing tasks. Hanna is observed using the parallel corpus tool or other resources to search out answers, but never analyzing the search results or the materials accessed through the corpus. Generally, Kate performed quite similar to Hanna and chose not to use the parallel corpus as instructed as a mediational tool for conceptual development, except for on exam 2. In completing this exam, Kate applied the parallel corpus as instructed in searching terms. She searched out words she knew, analyzed the concordances and then composed sentences (see Figures 4.8 and 4.9). This resulted in the most complex and detailed sentences Kate produced all semester, notably using the terms she had searched for and analyzed in order to create dynamic sentences for the exam. She used the knowledge she already had, as she only looked up terms she knew (evidenced by her searching in Chinese), but needed further mediation in grammatical and functional construction – which the parallel corpus provided. On this one exam, Kate used the parallel corpus in a way that led her conceptual development of the terms 非常 (*extremely*) and 一起

154
(together) as evidenced by the sentence she constructed and meaning she conveyed using these two terms. Despite the success on this one exam, though, Kate returned to her prior approach to learning and focused on searching out answers instead of developing understanding.

Cathy, who also belonged to the category of learners who were resistant to the parallel corpus, is a unique case. Cathy evidenced linguistic and conceptual development throughout the course of the study. However, with the exception of the zodiac story assignment, she rarely turned to the parallel corpus for any mediation, choosing instead to pursue traditional learning resources to understand the material and concepts. From the first month of the study she was observed developing and internalizing concepts such as 漂亮 (beautiful) before the parallel corpus was introduced, and her approach to learning remained unchanged through the course of the study. She did well and evidenced further development of the language during the second and third months of the study, but nothing links her development to the parallel corpus. While it was never hypothesized that learner development was impossible without using the parallel corpus, the assertion is that the parallel corpus would enhance learner experience and capacity for development. Thus in Cathy’s case, the question remains as to whether she would have been able to attain a higher level of development had she incorporated the parallel corpus. Contrasting her experience with Ellen, who was quite similar to Cathy other than her choice to actively incorporate the parallel corpus, suggests that Cathy could have progressed much further had she chosen not to be indifferent toward the parallel corpus and its application.
5.1.3. Contrary to Development

The final individual categorized as being resistant to the parallel corpus was Isabel, who also proved to be a unique example in her application of the parallel corpus. She went from using the corpus in a way that was either indifferent or counterproductive to her language development, to fully embracing the technology and demonstrating great progress. Her initial use of the parallel corpus, though, proved to be more harmful than good. In completing the zodiac story assignment she continually had all three language formats open, but focused primarily on English. While this same behavior was observed in Frank and considered indifferent, Isabel’s use differed in that she demonstrated little analysis of the Chinese terms. She would search out terms in English, and then guess what the correct corresponding Chinese term was. During this assignment as well as on the exam she searched for terms, copied what she found, but clearly did not understand what she was writing. The parallel corpus helped her to form an answer, but not to understand the term or concepts she was using.

At different times, all four boys from the poor performers category were also observed using the tool the same as Isabel to search for answers and simply guess. These instances are counterproductive to development as the corpus evidently allowed the participants to complete tasks without understanding the material. An online dictionary may likely have required greater cognitive interaction and understanding than using the parallel corpus to search in English and simply guess at the corresponding meaning. Fortunately this behavior was minimally observed among the boys in the poor performers category, and Isabel evidenced a major transition in her approach to learning by the end of the second month of the study as she recognized her previous approach was not
leading to the desired results both in terms of her grade and her development. As will be discussed in greater detail in the following section, in the final month of the study Isabel used the parallel corpus to model her sentences after, and focused on the Chinese format in the corpus – making the effort to understand ideas and concepts through Chinese before then attempting to use the terms and grammar to create her own meaning. This transition clearly demonstrates that the corpus allowed her to conceptualize the term, and then create her own meaning based on her new understanding.

5.1.4. Evidence of Development

When participants were observed both participating in learning activities and using the parallel corpus as instructed, various levels of constructive achievement and development were observed. Several studies have already clearly demonstrated that parallel corpora can function as a tool that mediates the researcher in disambiguating linguistic variation and developing conceptual understanding (Johannson 1999; Johannson & Hofland 2000; Ebeling 1998). The current study now demonstrates how a parallel corpus functions to mediate learners’ conceptual development and language ability.

Diane provides a clear example of this mediated development. From the beginning of the study Diane was observed to be a diligent learner who focused on understanding the language and new material. In the first month of the study she was observed struggling with the concept of classifiers on one of her assignments. By the exam, however, she clearly demonstrated that she had not only become familiar with several classifiers, but could use them correctly in forming unique sentences (section 4.1.3.1). This evidenced a conceptual understanding of classifiers as demonstrated by her
ability to correctly use them in several instances. While Diane clearly evinced the ability and desire to understand linguistic constructs on her own from the beginning, the question then becomes how the parallel corpus was able to aid her in this development after it was introduced. As Arievitch and Stetsenko (2000) posit, leaving a student as Diane to develop her own strategies leads to less productive performance, whereas providing mediation, or a tool for mediation, provides students with the necessary conceptualization for more efficient learning.

This assertion becomes apparent in Diane’s first experience using the parallel corpus. Initially she used the tool much the same as the individuals in the poor performers category, and opened all three language formats and focused primarily on the English text. Approaching the task in this way she appeared to struggle with the assignment and move slowly through the story. About twenty minutes into the screen recording, she is then observed completely adjusting her approach to the task and use of the parallel corpus tool, by following the directions of use provided by the instructor. She closed the additional language formats, focused primarily on the Chinese, and then began to navigate the story and the rest of the task far more efficiently. In Diane’s case it took this first twenty minutes to struggle with learning to use the parallel corpus as a mediational tool, but once she developed this understanding she was observed using the corpus in this manner the remainder of the semester.

Diane’s continued language development was also apparent during the final month of the study as she demonstrated the ability to not only use the material discussed in class to develop dynamic sentences and writings, but to also incorporate additional constructs such as the term 都 (all/both), which had not been discussed in class.
Additionally, she was observed to be more efficient in her work by being able to demonstrate this level of mastery in less time. For instance, she was able to complete the final exam in half the allotted time while others struggled to complete it at all. This is also a marked improvement for Diane herself in comparison to performance in prior months. The last key indication to note is that Diane self-reported using primarily the parallel corpus and notes in completing work and assignments. This self-reporting becomes a key indication of the parallel corpus’ successful mediation in Diane’s development, as very minimal use of the resource (or any other online resource) was observed in Diane’s screen recordings. By the time of exams, Diane had internalized concepts and required little additional mediation to complete the tasks and convey her meaning using the new material and constructs.

Through considering these several factors of Diane’s performance, a strong implication emerges suggesting that the parallel corpus played a key role in mediating Diane’s improved development. Though in Diane’s case it is difficult to definitively make this correlation, all considerations seem to indicate it. A quote from Diane’s end of semester questionnaire further substantiates these claims, as she stated that she appreciated being able to analyze several example sentences and “the translation, pinyin, and character for all the words [she] needed”. This statement implies that Diane used the parallel corpus as a tool to develop functional understanding of the “words” and concepts she looked up, and as evidenced in her written work she was then able to apply these terms in new contexts. This ability to re-contextualize concepts is a key evidence of language development according SCT research (Ferreira & Lantolf 2008), and unlike discussed previously in Frank’s example, Diane not only observed the re-
contextualization of concepts by using the parallel corpus, but she was then able to re-contextualize these constructs herself, clearly demonstrating how the parallel corpus had aided her understanding and language development.

Gail’s performance during the study also provides evidence of how the parallel corpus aided her development, with a couple of very clear examples demonstrating how the corpus was used and the conceptual understanding that resulted. Baobao, Danielsson, and Teubert (2005) observed that the structure of parallel corpora often leads learners to analyze texts in multi-word conceptual units as opposed to individual lexical units (p. 132). This finding further asserts that parallel corpora mediate learning by guiding learners to analyze text in terms of meaning within context rather than just considering lexicon or grammar. Gail’s experience in learning the term 姓 (surname) provides a very clear example of this conceptual understanding and development (section 4.1.3.3.). In her learner log she specifically reported using the parallel corpus to search and analyze this term. Her performances on an assignment as well as an exam using this term as both a verb and a noun clearly depict not only her conceptual understanding of it, but that the parallel corpus guided her learning in a way that led to her understanding this term as a “conceptual unit”. Several parallel corpus scholars (Pereira 1996; Peters, et.al. 2000; Tsui 1996) have suggested the potential of parallel corpora in functioning to aid learner development in this way, and Gail’s example provides a clear validation of their researched suggestions through her mediated development of this term as well as other constructs.

Both Diane and Gail’s examples are of specific lexical items and illustrate how the parallel corpus functioned to mediate their development. At a basic level, their
acquisition of lexical items is congruent with several other parallel corpus studies. Tsai and Choi (2005), for example, noted significantly improved lexical development among students using parallel corpora, also remarking that students were able to recognize multiple functions of polysemous terms. Wang (2002) demonstrated how a parallel corpus assisted learners in understanding the grammatical variation in how the term 现在 (now) functions in Chinese versus English. Frankenberg-Garcia (2000) described how Portuguese L1 learners of English were able to acquire a functional knowledge of how the English preposition with aligns with the Portuguese com, even though the two words do not perfectly overlap in meaning and use. These three studies present different aspects of development associated with specific lexical items: meaning, grammatical structure, function, and use. By analyzing both the process and product of Diane and Gail’s use of the parallel corpus, it becomes apparent that their acquisition of a “lexical item,” actually went far beyond simply learning a new term. They were able to analyze the new term and exhibit development in all four of these aspects, clearly demonstrating an understanding of meaning, function, and use.

The observation of this dynamic conceptual development of these specific items becomes clear when analyzing the complete picture through SCT. Parallel corpora enable learners to compare lexical phrases in several ways. They not only allow the learner to observe multiple examples of the use of a term, but to then compare and relate those uses to similar patterns in both their L1 and the target language. This allows the language learner to make direct connections between how such concepts in the second language relate to the familiar concepts of their first language. Undoubtedly, being presented with numerous concepts when searching out a single term initially complicates
the learning process, the ultimate outcome leads to a greater level of understanding of the term. As SCT stresses, development requires that an individual’s encounter with a principle “generate[s] a crises which the individual must overcome and in doing so, he or she develops” (Lantolf & Thorne 2006, p. 313). In Diane and Gail’s examples they initially were searching for the meaning of a term, 都 (all/both) and 姓 (surname) respectively. Instead of simply finding a word for word correspondence and moving on as would be done when using a dictionary, using the parallel corpus required them to analyze several concordances, observe how the term was used in Chinese and how it was represented in the aligned English translation, and ultimately develop an understanding not just of a lexical definition, but of function and use. As stated, this approach initially complicated the learning experience, but resulted in greater conceptual development and understanding.

Gail and Diane were not alone in using the parallel corpus in this way, as both Leo and Ellen exhibited similar patterns of corpus use and conceptual development. All four of these participants were independently observed using the parallel corpus to search for new vocabulary items, analyze the concordances to derive meaning and use, and ultimately re-contextualize the terms by creating their own sentences and writings incorporating the queried terms. Aijmer (2008) states that parallel corpora are useful to “disambiguate polysemous items, reveal the degree of mutual correspondence of lexical items in different languages, and uncover cross-linguistic sets of translation equivalents in the languages compared” (p. 99). The above mentioned participants were observed using the parallel corpus in all of these ways, but also then going on to demonstrate their acquired knowledge through written performance on exams and assignments.
Leo was unique in how he used the parallel corpus in that he tended to explore content beyond the material being discussed in the classroom. While other participants were observed using the parallel corpus primarily to search for a few terms related to the content being discussed, or to affirm the correct understanding and function of a term they were already familiar with, Leo regularly explored the parallel corpus for additional content often then creating original, humorous, yet functionally correct and dynamic sentences and writings. A key observation of his methodology was that anytime he searched for a new term, he was then observed analyzing the sentences and texts in which the term was found. Often times this would result in him spending a significant amount of time reading through portions of several source texts to attain a better understanding of the term and context in which it was being used, and undoubtedly boosting his comprehension and development. As Frankenberg-Garcia (2005) points out “When reading in a foreign language, L2-L1 parallel concordances can help learners to understand foreign words, meanings and grammar that they are unfamiliar with. Extracting concentrated examples of parts of foreign language that they don’t quite understand, matched to equivalent forms in their mother tongue can help boost language comprehension” (p. 194). As Leo took time to read through additional texts he was exposed to a deeper contextual and conceptual understanding of the meaning and use of the terms he was searching for, as well as being exposed to many additional characteristics and features of the language.

This approach to language learning is unique to parallel corpora, as being able to compare the Chinese and English texts enabled Leo to rely on his L1 to develop a conceptual understanding of what was being expressed in Chinese. In others words, the
parallel corpus afforded him the opportunity to explore additional ideas that would have likely been beyond his comprehension through other learning mediums. Fan and Xu (2002) made similar observations with regard to increased reading comprehension, and noted that once students became familiar with analyzing texts through parallel corpora, they tended to prefer this medium above all others compared. Leo’s performance was consistent with the students in Fan and Xu’s study, as he exhibited a strong preference for the parallel corpus, using it almost exclusively as a learning resource after being introduced to it. In terms of development, Leo was observed not only using the parallel corpus to improve his conceptual knowledge of the language to complete course tasks, but also actively seeking out new material and content to explore on his own. In essence, the parallel corpus functioned to mediate Leo’s understanding of confusing lexical items and linguistic concepts, but it also helped him develop a new and very productive approach to learning how to learn the language.

The two remaining participants to be discussed, Isabel and Ellen, serve as perhaps the two most explicit examples of how the parallel corpus functioned to mediate development and allow these learners to perform beyond their capacity without the corpus. Isabel provides a contrastive illustration transitioning from the second month of the study where she exhibited use of the parallel corpus tool contrary to development (section 5.1.3.) to then changing her learning approach and making great strides in her learning. Ellen provides an idealistic example of how a top student’s performance can be greatly enhanced by the appropriation of parallel corpus technology.

Mauranen (2002) noted that an individual’s first experience with parallel corpora is generally marked by observance of a much larger range of translation equivalents than
encountered in any dictionary. This distinction between these two types of resources seems to be what caused Isabel to struggle with implementing the parallel corpus initially, as she appeared to use it similar to how one would use a bilingual dictionary. After her counterproductive application of the parallel corpus during month 2, the instructor was able to meet with and mediate her understanding of how to successfully use the tool to aid her learning. During the third month of the study Isabel began to demonstrate great improvement. While she still struggled with creating grammatically correct sentences, it became clear that she both understood the content of what she was reading and writing, and that she was able for the first time to express the meaning she wanted to through the language. In prior months her written responses clearly emulated the textual examples with little to no variation, often times clearly demonstrating that she did not understand the meaning or function of the terms and constructs she was using.

In the final month of the study once Isabel began to use the parallel corpus as directed, her language learning experience clearly transformed. The sentences and writings she composed were far more complex and dynamic, and through screen recordings and learner logs it was apparent that the parallel corpus played a key role in this transformation. For the first time Isabel was analyzing the texts and developing a conceptual understanding of how the language functioned, and she was able to then demonstrate this conceptual knowledge in her written responses. Karpov (2005) argues that concept-based instruction affects “the cognitive development of students such that when they approach new tasks and problems they follow a theoretical and systematic orientation” (p. 189). Isabel’s example is quite illustrative of this claim, as once she understood how the parallel corpus could function to mediate her development, she
completely altered her approach to the remaining tasks and demonstrated a systematic orientation which led to much better performances that indicated her conceptual understanding of the language and constructs she was using. In accordance with SCT as noted by Karpov (2005) and Lantolf (2008), the conceptual approach to learning enabled Isabel to develop a cognitive theoretical knowledge that she was able to apply in subsequent learning tasks, and will hopefully continue to apply in her study of Chinese.

Ellen provides an illustration of how the parallel corpus was able to take an exceptional student to new levels of achievement and development. From the beginning of the study, Ellen was noticeably one of the top performing participants and evidenced a deeper understanding of material and concepts. She tended to construct more complex and dynamic sentences than her classmates, and regularly incorporated grammar constructs and lexical items from several previous lessons indicating that she had achieved a functional understanding of the material she had learned. Based on her performance, it was clear that she was already internalizing the material being covered in class, as she was able to clearly express her own meaning through the language. When the parallel corpus was introduced, Ellen used the tool as directed by the instructor and focused almost exclusively on the Chinese text using the other language formats and corpus features only as needed. On the questionnaire she even indicated that one of her favorite features of the tool was that “you could click on a character and it would tell you the pinyin but not the English. That way, if you knew the pinyin you could think through what the sentence meant”. This statement aligns with what was observed in her analysis in that she consistently focused on understanding the ideas being expressed in Chinese, and truly sought to work through the language. Ellen proved to be a model participant
who not only fully engaged in the curriculum, but also incorporated the parallel corpus as directed in order to achieve optimal development. It is apparent that Ellen did not rely on the parallel corpus the same way as many other participants, but instead used it to expand her understanding of the language beyond the limits of the lessons covered in class.

For example, on Ellen’s written assignments, she not only used basic new terms and structures similar to other participants, but she applied more specific and complex terms mentioned in the text. She was able to correctly derive meaning and understand the terms solely using the corpus. This observation indicates that Ellen’s ZPD was beyond that of most of her classmates, signifying that she was not challenged as much by the classroom content which focused primarily on challenging the classroom ZPD, or the average student in the class. Despite this distinction, though, Ellen’s application of the parallel corpus enabled her to continually be challenged beyond the material covered in class, so that she could also attain a greater level of development. Not only did she clearly master the content covered in each lesson, but by using the parallel corpus she also developed a conceptual understanding of many additional lexical items and structures that she encountered in the parallel corpus and then regularly applied in her written work (section 4.1.3.3.). Thus, the adaptation and incorporation of the parallel corpus notably accelerated Ellen’s level of development, by continually challenging her with new material and content.

While Ellen was consistently at the top of the class both before the study and during the first month, her language ability began to increase at a greater rate than her peers after she began learning with the parallel corpus. While her development cannot be attributed solely to the corpus tool, this resource was a key factor in the rapid rate
increase in her development and performance. By the end of the three month study Ellen had made such significant progress that it was clear that continuing to study with her current cohort would limit her individual growth. At the beginning of the next term Ellen was actually advanced to Chinese 3, skipping Chinese 2, and continued to thrive in this new setting.

While many factors contributed to Ellen’s overall success, the adaptation of the parallel corpus became a key factor in both her performance and her conceptualization of Chinese. Similar to Isabel, it altered her approach to learning, allowing her to develop conceptual knowledge she was able to apply in later learning tasks. Furthermore, it allowed her to consistently be challenged and perform at the top of her ZPD, instead of being restricted by activities focused on the classroom ZPD. The extent to which Ellen adapted the parallel corpus’ systemic orientation to learning Chinese is solidified in a look at the notes from her final presentation.

As stated in the data analysis of the final presentation (section 4.1.4.), the notes from participants’ final presentations were not initially envisioned as being part of the study or the data analysis as students were not required to have notes or other written material for the presentations. However, five participants did have note cards with their presentations scripts, and submitted this written material. As highlighted in the data analysis, Cathy and Diane’s notes were only in Chinese with no pinyin or English. As this was an oral task, the lack of pinyin stands out. Cathy and Diane made a conscious decision to use only Chinese characters, and to prepare their presentation scripts so that they necessarily were able to recognize and pronounce each character included in the script. As many of the words each participant used were new vocabulary items unique to
the topic of their presentation, they undoubtedly had to spend time learning the character, meaning and function of each term. The aspect that truly stands out is the focus each student places on characters. One of the guiding purposes of the parallel corpus was to focus on learning characters by omitting the obligatory pinyin present in most traditional beginning level Chinese learning resources. The pinyin was of course still accessible in the parallel corpus as needed, but the focus was placed on the characters. The final scripts composed by both Cathy and Diane strongly suggest that they had internalized this approach to learning Chinese, and therefore exerted the extra effort in learning the characters for their presentations.

The scripts submitted by Isabel, Gail, and Ellen also suggested that the parallel corpus had an effect on their approach to learning Chinese. All three of these participants prepared their scripts in the three language formats, Chinese characters, pinyin, and English, mimicking the structure and presentation of text in the parallel corpus. This observation is significant as it suggests that these individuals had begun to conceptualize the Chinese text through these formats. Thus in addition to developing conceptual knowledge of specific aspects of the Chinese language, they had also internalized the approach to interacting with written Chinese as afforded by the parallel corpus. Whether this is a result of them using the parallel corpus, or mere coincidence, it is evident that these three participants found it most useful to conceptualize the language in these three formats. Observing this final behavior, though, strongly indicates that the parallel corpus not only had a large effect on their language development (product), but in the way they approached learning the language (process).
5.2. Discussion of Pedagogical Implementation

The second question researched sought to explore how the students and instructor applied the parallel corpus throughout the semester. Aspects of this topic were discussed in the previous section as how participants adapted the corpus proved to directly affect their development. This section goes into greater detail in order to consider what specific practices of implementation by the participants as well as the instructor proved to be most beneficial, and what approaches should be altered and improved in future applications. The purpose of this discussion, therefore, is to provide direction on what practices best promote effective and efficient implementation of parallel corpora as mediational tools in foreign language classrooms.

5.2.1. Learner Implementation

As evident through the analysis of data and the discussion thus far, learners chose to implement the parallel corpus in many different ways. One key distinction linked to development was whether participants primarily concentrated on the text in their L1 (English) when they accessed the parallel corpus, or whether they focused on the target language text (Chinese). Those from the poor performers and resistant to parallel corpus categories were observed to focus first on the L1 text and refer to the Chinese mainly to search out specific answers. On the contrary, those who integrated the parallel corpus (and ultimately exhibited higher levels of achievement and development) chose to concentrate on understanding concepts through the target language when using the corpus. As detailed in section 5.3. Discussion of Learner Attitude/Perception, regardless of which approach participants exhibited, all participants found the parallel corpus helpful and useful in their language learning. Interestingly, these findings coincide with
the student responses to the questionnaire in Fan and Xu’s (2002) study. They also observed that a majority of participants relied primarily on the text in their L1, a minority relied mainly on the L2 text, but 85% considered using the parallel corpus more useful than other resources. These combined findings suggest that regardless of approach, implementing the parallel corpus is perceived as being valuable by learners, and actually preferred over other mediums. While Fan and Xu’s study concluded with this observation, the current study goes on to consider the effectiveness of these different approaches as linked to development and learning outcomes.

Based on learning outcomes it is evident that those who focused on the target language when using the parallel corpus performed better and had greater levels of development. This realization is of importance for two primary reasons. First, if focusing on the target language text is more effective, then pedagogical applications of parallel corpora should obviously emphasize using the corpora in this way. Participants in the current study were in fact instructed to focus on the character format and use the other formats as necessary to supplement their understanding. Despite this instruction, the majority still went on to focus primarily on the L1 text. This leads to the second point, that student use of parallel corpus should be mediated more directly to ensure optimal implementation by focusing on the target language text. The implementation of the parallel corpus in the current study fell short in this regard, but analysis highlights how some of the design features of the parallel corpus did mediate learner development well, while other mediational features are lacking. This additional mediation could be accomplished by the instructor, by adding additional features to the parallel corpus interface, or by using different practices when using the corpus.
Considering which approach to take is a complicated decision, as the success ultimately depends both on the feature of the corpus as well as the user. This is illustrated by the contrastive experiences of Adam and Diane. They both used the pinyin search feature of the parallel corpus in the same way to search the exact same term, but with two different learning outcomes. When searching pinyin, the parallel corpus required the user to identify the correct tone mark in order to complete the search – thus requiring a specific path of implementation to use this feature. The difference then came as Adam chose to quickly find the character and move on, while Diane analyzed the search results (focusing on the target language) to develop an understanding of the term and its use – Adam’s approach leading to a quick answer, and Diane’s approach leading to development.

In this example, both participants were afforded the same learning opportunity but chose different paths. Laviosa (2002) identified a significant finding that parallel corpora are able to reveal the precise information “the learner needs to acquire in order to establish mental links between L1 and L2 schemas and create new L2 schemas when there is not reciprocity between the two languages” (p. 110). Adam and Diane were both provided with this precise information, but used it differently. From an SCT perspective, this application of the parallel corpus offered both participants an opportunity to develop through mediation to achieve optimal development, but only Diane did. Unfortunately, Diane belonged to the class minority in optimally implementing the tool in this instance and others, leading to the question of what aspects stood in the way of others using the tool in this way and exhibiting similar positive development.

Two parallel corpus scholars provide insight into this question. Aijmer (2009)
points out, “Teachers (and learners) look for simple answers to grammatical problems in terms of what is right and wrong and shy away from the fuzzy picture of language as used in the corpus concordance” (p. 2). Even though it has been demonstrated in the current study and others that corpora can provide a more complete conceptual framework for understanding linguistic phenomenon (see also Bernardini 1996), the potential complexity of corpora could easily discourage a teacher from opting to use this technology, and as has been observed, dissuade learners from using them as instructed. In other words, it is a new technology and new approach that requires additional effort not just on the part of the instructor, but also on the part of the student who must choose to make the effort required to effectively adapt and learn through the parallel corpus. Thus a main consideration that should always be at the forefront when incorporating parallel corpora in the classroom, is not just that participants are using the resource as guided, but that the instructor has a way of observing and continually guiding participants in using the resource.

The first experience participants had in using the parallel corpus, the zodiac story assignment, observed many learners using the corpus as directed, and continued mediation from the instructor in using the resource. The instructor noted that it was a positive experience for all and that “students seemed to enjoy the activity and were quick to point out that there were many words and characters that were new to them – but that they were able to figure out meanings all the same” (section 4.3.2.). From screen recordings we know that not all participants used the tool as directed even in this initial activity, but that it was a positive learning experience and the instructor was actively mediating and guiding participants as they became familiar with the parallel corpus.
Following this first experience, participants continued to use the corpus both in and outside of class, but instructor mediation of their use of the tool was not as direct. At this point in the study most of the participants readily adapted the parallel corpus and continued to use it, but did so in a way that complemented their existent learning approaches. In some instances (Diane, Gail, Leo, Ellen) this led to positive outcomes as these participants’ approaches to learning focused on understanding the content and developing the language. In Cathy’s case her language development continued, but she was chose not to use the corpus. In most other cases (Adam, Brian, Frank, Jack, Hanna, Kate) participants used the parallel corpus to search out answers, but not fully develop conceptual knowledge of the language. While this practice was consistent with their previous approach to learning, the implementation of the parallel corpus could have become a more transformational activity in assisting at least some of these participants in learning how to learn the language. This is a challenging task, but one that prompts innovation in continued adaptations of parallel corpora as well as other learning resources.

Important to point out is that even though some participants chose to use the tool contrary to direction, they were aware of its intended purpose. Hanna, for example, was only ever observed using the parallel corpus to search out answers, and her screen recordings and performance provided little evidence of how the parallel corpus aided her development. Despite this, she stated on the end of semester questionnaire “nciku [online bilingual dictionary] helped me find what new vocab items meant, but parallel corpus helped me figure out how to use them” (section 4.2.2.1.). This statement suggests that she either used the parallel corpus tool correctly at home, and was simply never observed
using it as directed in the classroom or the data, or that she understood and repeated the intended function of the parallel corpus in her response while choosing not to actually use it in this way. Either way, her example illustrates that when implementing a new learning resource, more than proper instruction is needed. Continued follow-up and mediation is requisite, but ultimately it is the learners choice in whether they choose to learn and what effort they are willing to commit.

Again returning to the examples of Isabel and Ellen this becomes further evident. As Isabel initially implemented the parallel corpus in a way that was counterproductive to her development, she required further mediation in order to change her approach and be more effective in her language study. In meeting individually with Isabel at the end of month 2 of the study, the instructor reinforced the directions that had been given in class, and then gave Isabel the opportunity to exhibit these behaviors with his supervision. On subsequent tasks, she was also given the opportunity to correct mistakes on assignments and receive additional feedback from the instructor. This adapted approach helped Isabel transition from concentrating primarily on her score in completing tasks, to instead focusing on understanding the concepts and ideas she was struggling with. This change in focus surfaced in part as Isabel realized her previous approach to learning was not working, but also as the instructor emphasized conceptual development and placed less stress on scores by allowing Isabel to repeat tasks on assignments as long as it was evident she was developing her understanding of the content. As a result, Isabel’s performance improved, but more importantly her conceptual development began to increase. The primary factor leading to this change being the way in which she implemented the parallel corpus, transitioning from focusing on answer seeking through
English, to understanding content through Chinese.

Ellen’s implementation of the tool focused on this latter approach clear from her initial experience in using it. Ellen’s example is extreme in that her adaptation of the parallel corpus required more effort than anyone else’s, but it also resulted in a better learning outcome than the others. Individuals such as Diane, Gail, Leo, and even Isabel by the end of the study, provide examples of what was anticipated in the research: participants using the parallel corpus to develop their conceptual understanding more efficiently. In essence, it was hypothesized that the parallel corpus would enable participants to achieve the same level of development as other resources, but with less effort, as the parallel corpus would make the learning process simpler and more efficient. This proved to be the case with all four of the individuals mentioned. Ellen, on the other hand, chose to put forth just as much or more effort as other learning methods would require, and consequently exhibited exceptional developmental results. In terms of implementation, Ellen’s example highlights the important distinction that technological learning resources such as parallel corpora should not be adapted simply to make the learning process more efficient, but to challenge students in a different way, leading to a greater level of development.

5.2.2. Instructor Implementation

A primary issue for an instructor to be cognizant of in implementing any learning resource, or simply preparing any lesson plan, is class size. The participants in this study came from two relatively small classes of nine and three students each. The smaller class size certainly led to a greater level of instructor mediation and individual attention than would be afforded in larger classrooms. Many of the participants explicitly recognized
that the smaller class size benefited their learning experience. While class size does affect learner experience and the implementation of new resources, it does not restrict the potential for using parallel corpora in larger class setting. However, it should be one of the initial factors considered in preparing to implement a resource such as a parallel corpus.

While class size is an underlying issue to consider, perhaps the greatest factor to recognize is how the adaptation of a parallel corpus changes the role of the instructor. Aston (1997) argues that corpora reduce teachers’ dependency on textbooks, and allow them “to concentrate on their role as learning rather than language experts” (p. 63). From an SCT perspective this is a major benefit of using parallel corpora as the teacher shifts from being a conveyor of knowledge to being an expert ‘co-learner’ helping to guide students as they develop. This change in roles is evidenced in the current study as participants who integrated the parallel corpus regularly turned to the tool to develop conceptual understanding of the language. The teacher was no longer the source for this knowledge, but a resource to assist students in analyzing the natural language texts in the parallel corpora. Accordingly, a discussion of instructor implementation of the parallel corpus becomes primarily consideration of how the tool was introduced and how the instructor adapted to mediate learning through the corpus.

As detailed in the analysis of the instructor autoethnography, the implementation of the parallel corpus into the existing curriculum did require a substantial amount of effort from the instructor. The process in setting up the corpus by adding course material and supplementary resources was labor intensive and typically required one to two hours of preparation per class period. Thus it can be observed that the introduction of this new
resource required substantial effort in preparing materials and adjusting lesson plans. Even during the preparation process, though, the instructor noted the perceived potential for pronunciation exercises using the corpus to review tone marks, instilment of content covered in class such as animal vocabulary, and exposure to real language examples of constructs covered in the class. These observations demonstrate that the instructor saw the potential value of the parallel corpus before it was incorporated in the class, but more importantly they illustrate how the instructor was not simply preparing to use the parallel corpus, but had specific topics and lessons in mind of how the parallel corpus could be used.

The initial introduction of the parallel corpus was designed around the zodiac story assignment. The purpose of the assignment was to help participants realize how the parallel corpus could be used to navigate Chinese text and concepts, as well as to practice and review content and vocabulary from the lesson on animals. Accordingly, the two main objectives were recognized as: 1) becoming familiar with the parallel corpus and its features, and 2) reviewing and internalizing the content from the lesson on animals. As has been observed, both of these objectives were achieved, and students reported positive learning experiences. One issue to draw attention to, which is considered in greater detail in section 5.3. Discussion of Learner Attitude/ Perception, is that while participants met the above two objectives and reported positive experiences, there was also an observed hesitancy toward the learning approach experienced in this exercise. Participants were overwhelmed by their first interaction with a complete text in Chinese. By the end of the activity many reported being encouraged by their ability to work through the text and understand it, but in general it was still a challenging task for the class. In terms of
implementation, it might be more effective to either introduce students to challenging conceptual learning activities such as this one prior to the introduction of the parallel corpus. While not necessary, it would likely be beneficial, as it would serve to prevent students from creating negative associations between the difficulty of the assignment and the use of the parallel corpus. In the study it is apparent that most participants were able to grasp how the parallel corpus was able to ease and mediate their performance on the task, but this realization would have likely been more apparent and encouraging to participants if had they struggled with a similar assignment without having the benefit of the corpus.

Beyond the initial implementation of the tool, the instructor continued to provide support and mediation throughout the rest of the study. He regularly used the parallel corpus himself in class while providing instruction, and frequently guided participants in their use of the tool on in-class activities. The instructor also noted more than once that students regularly struggled with understanding grammatical constructs by themselves through analyzing texts in the parallel corpus, and often required additional instruction and mediation. In these instances, the instructor mediated through the parallel corpus by directing participants in their use of it and then further explaining constructs using the texts and concordances in the corpus. In doing so it was also noted that this additional mediation was typically focused on both instructing the content of the term or content under question, and on guiding participants in how to resolve such issues independently in the future by using the corpus.

Beyond the issues already mentioned, there were a few specific examples identified where the parallel corpus could have been implemented more effectively.
Isabel was a great example of how an altered approach in using the parallel corpus took a student from being counterproductive, to then making great improvement in developing language constructs. While Isabel’s counterproductive use of the parallel corpus was apparent during the second month of the study after analyzing her performance, it was not until the end of this month that she received the necessary instructor mediation to change her approach and begin to make great improvement. While several factors contributed to this learning process, an earlier intervention by the instructor would have likely better-facilitated Isabel’s development.

Similarly, it was not until Isabel’s individual meeting with the instructor that he began to more effectively guide her in her work, allowing her to redo it if necessary. Once the instructor made this change to his approach, Isabel’s focus transitioned to being on content and concepts instead of scores. A comparable example is also found with Ellen, who was marked down on exams for mistakes she made on exams, even though those mistakes resulted from her engaging in more complex constructs and writing more dynamic responses than the exam prompt required. These observations suggest that while the instructor had made a clear effort to guide students in learning conceptually, he overlooked aspects of his grading and feedback that still placed emphasis on score instead of development.

Next, while the exams and materials were designed in a way to challenge students by requiring them to compose writing and create their own meaning with the language, some participants such as Hanna worked around this. Instead of using the corpus to aid in her meaning making, Hanna managed to use it and other resources to quickly search out answers. Again this returns to the challenging task of guiding the learner’s desire and
approach to learning, but a careful exam design could help minimize such occurrences in the future. The final issue to restate is the challenge of glitches when using technology in the classroom. Not only did the instructor have to put extra effort into preparing the digital resources for the parallel corpus, but also in preparing alternative plans in case technological issues arose on days in which the parallel corpus was to be used in class. This lesson was learned the hard way as the server crashed the first day the instructor intended to use the corpus, but proved to be a cautionary lesson the instructor learned early on, and that should obviously be considered by others as they seek to implement parallel corpora and other technology in the classroom.

5.3. Discussion of Learner Attitude/Perception

The final research topic to be addressed is learner attitude and perception. The study sought to understand how students viewed the application and implementation of the parallel corpus tool. Generally, this translates into reviewing students reported perception and experiences. More specifically, it also includes a discussion of how students reported attitude aligns with their observed performance. The purpose in addressing this questions is to both understand the aspects of the parallel corpus that were well received, and also to consider how the pedagogical implementation of parallel corpus technology could be improved in future applications.

Participant feedback and reported experience with the parallel corpus was overall very positive, indicating that the corpus was well received and appreciated by participants. As previously indicated (section 4.3.), there was little to no resistance to using the parallel corpus, simply varied levels of successful application. The participants were overall eager to use the new resource and explore how it would assist them in their
language learning. This observation is undoubtedly unique to the population of students participating in this study, as all of them were familiar with and comfortable using technology. Among other populations with lesser experience with technology and digital resources, a greater level of resistance would likely be anticipated. However, for the participants in this study the parallel corpus was not simply well received, but in fact encouraging to their language learning experience. Exploring a new language learning technology proved motivating and ultimately enhanced participants’ engagement with the course content.

The likert analysis completed by participants on the end of semester questionnaire suggests that participants found the tool very useful and planned on using it in future settings to continue their study of Chinese. As has already been acknowledged, the statistical validity of the likert analysis is lacking due to the small population size, but it is indicative of learner attitude, and what was observed in the qualitative analysis. Furthermore, the apparent lack of criticism offered by participants in both learner logs and the questionnaire substantiates the general observation of a positive perception of the parallel corpus. As noted in the analysis, there were three primary suggestions offered for improving the parallel corpus, but even these critiques were offered as ideas for enhancing a resource that was already perceived as useful. Specifically, a few participants highlighted the formatting glitch that “scrunched” text together when all three language formats were opened, several students also mentioned they felt the tool should have a word for word translation feature, and Kate suggested making it possible to search terms in all language formats without having to specify which was being used.
The first issue is a programming issue that undoubtedly needs to be addressed and corrected to improve user experience with the tool. The latter two critiques, however, identify intentional design aspects intended to promote development, and not necessarily ease user experience. The incorporation of a feature that offered word for word translation was intentionally excluded, as this would have negated one of the primary conceptual learning features of the parallel corpus. By not providing word for word equivalences, participants are pressed to engage with complete texts in order to understand the meaning (usually through the L1) and then observe how that meaning is expressed in the L2. As Aston (1997) stresses, this is one of the primary features of corpora that enhances language learning, as students are able both observe and analyze “regularities of language use” (p. 63). It is not surprising that several students would prefer having word for word translations, as this would have provided an easier method for quickly identifying words. Excluding this feature, though, required the students to struggle with the meaning of the texts in order to derive both what was being said, and how that meaning was expressed in the target language. As Johnson and Golombek (2011) point out, engaging in a cognitive struggle such as this is necessary for conceptual development.

The critique offered by Kate, that the parallel corpus should not require users to specify language format when searching, is a valid point that was considered heavily in designing the tool. While making it possible to search without specifying language format would be simpler, the program would then not be able to decipher between English input and pinyin input as both use the Latin alphabet. It was necessary for these two formats to be distinguished so that the user would be forced to indicate the correct
tone mark above vowels when inputting pinyin. This design feature was incorporated in order to promote students’ learning of tone marks by requiring that they be correctly indicated. As demonstrated in the analysis, both Diane and Adam used this feature in order to more efficiently search for terms in pinyin, and found it very helpful. While Kate’s suggestion would simplify the search feature, it was ultimately concluded that requiring the user to indicate tone mark when searching in pinyin was more beneficial for learner development. In the future, further options could be explored to incorporate Kate’s suggestions without having to sacrifice the required tone mark indication when searching in pinyin.

With the exception of these few critiques, participants’ comments and reflections at the end of the study indicated positive and encouraging attitudes toward the parallel corpus. As Leo indicated, he felt that parallel corpus offered him everything he needed to complete the course tasks and develop his language ability. While this perspective is very encouraging, it was obviously not the only one. Initial reactions to using the parallel corpus tool on the zodiac story assignment evidenced some hesitancy among users. The expressed hesitancy was not toward the parallel corpus itself, but rather toward the approach of learning through complete concepts and working through an entire text that had several new vocabulary items. Reading the zodiac story was the first experience participants had in reading anything longer than the brief one to two sentence responses in the textbook dialogs. Thus the task of working through an entire text and being able to grasp the thesis as well as the specific concepts expressed was challenging for students. It might have been more effective to expose learners to complete texts in traditional formats before introducing the parallel corpus, so that they could then experience the
contrast between the two approaches to analyzing texts in the target language, and not create an association between the parallel corpus and the challenging learning activity. Participants reported experience on the zodiac story assignment was very positive, and there was a lot of positive feedback on the parallel corpus. However, after this initial assignment several users such as Cathy did not continue to use the parallel corpus on future tasks. It is possible that participants associated the parallel corpus with the challenges they faced in completing the zodiac story assignment, and this association could have discouraged them from turning to the corpus tool in future applications.

While there is no firm evidence to derive this conclusion, the logical possibility does suggest that future first applications of parallel corpora in pedagogy might experience greater success by planning an initial activity that is less challenging, but that clearly demonstrates the unique benefits of this technology. Although, participants in this study could have had a better initial experience, there experience was still positive. The challenging nature of the task forced them to focus on the meaning of the text and on understating what was being expressed through the language instead of learning a list of terms and constructs. While this was a large transition in learning, participants were able to identify the value of the parallel corpus. This is evident in the analysis of the instructor’s autoethnography, as he observed that “the students seemed to enjoy the activity and were quick to point out that there were many words and characters that were new to them – but that they were able to figure out meanings all the same” (section 4.3.2.). Thus while the assignment was challenging, and perhaps more challenging than would be ideal for students’ first experience using parallel corpora, they were able to
successfully complete the task and begin to see how the parallel corpus could assist them in understanding concepts and ideas through the target language.

As participants experience with the tool evolved, it became apparent that they valued the design of the parallel corpus with multiple language formats. Interestingly, Cathy mentioned during the first month of the study (before the corpus was introduced), that she desired a language learning resource that allowed her to click on characters to see tone mark and pinyin. This desire was realized once the parallel corpus was introduced, and students repeatedly expressed how much they liked having access to tone mark and pinyin, either by selecting language categories, or by the interactive feature that displayed these items when a character was clicked on. As demonstrated in the analysis (section 4.2.4.2.), Cathy, Gail, and Isabel found it very useful to have all language formats in one screen, and Jack identified the “easy usability” of the tool in navigating the language formats as being very appealing. All of these comments clearly indicate that participants were overall happy with the parallel corpus, its design, and their experience using it to develop the language.

A key aspect of user experience and perception is directly related to corpus design. Lavid, Hida, and Zamorano-Mansilla (2010) illustrated the superiority of using a parallel corpus designed for pedagogy when teaching students, as opposed to one used primarily for linguistics research. Their observation is further substantiated by the positive experiences of the participants in this study, as individuals were not hindered by the complexity of navigating additional linguistic research features that were beyond the realm of their language learning objectives. The parallel corpus used in this study was created specifically for pedagogical application, and the participants were able to focus
on learning the language, and highlighted several design features that improved their experience. Leo, Diane, and Ellen all identified different aspects of the parallel corpus design that they found desirable and very useful (section 4.2.2.2.). While a few participants did not apply the parallel corpus to its full potential and evidenced a preference for traditional learning resources, a majority explicitly expressed and evidenced preference for the parallel corpus.

This finding is consistent with Frankenberg-Garcia’s (2005) observation that when presented with multiple learning resources for reading comprehension tasks, students preferred using the parallel corpus. The majority of participants in the current study who actively engaged in the course evidenced an apparent preference for the parallel corpus both in their learner logs and the questionnaire, as well as in their observed performance. Frankenberg-Garcia reasons that students exhibit this preference in part because parallel concordances enable language learners to boost comprehension and gain conceptual access to texts that might otherwise be beyond their level, or as St. John (2001) states, parallel corpora help students develop the ability to learn how to learn. Ellen, Leo, and Diane explicitly uphold this reasoning. As previously indicated, Diane stated that she loved using the parallel corpus to read stories in Chinese as it enabled her to not only read the texts but to understand them (section 4.2.4.1.). This statement not only substantiates Frankenberg-Garcia and St. John’s claims, but also clearly demonstrates how several participants came to value the parallel corpus as a resource that allowed them to access meaning through Chinese.

Perhaps this point is made most clearly in the instructor’s autoethnography as he observed on the second to last day of the course that three separate students individually
approached him and requested a specific story be uploaded to the parallel corpus so they could read through it over the summer break. These requests clearly indicate participants’ preference for using the parallel corpus tool for reading, but also suggest that participants felt the parallel corpus functioned as a resource that allowed them to access material that would otherwise be beyond their conceptual reach. As discussed in learner achievement, when applied appropriately the parallel corpus is able to mediate user development and stretch them to work at the top of their ZPD to develop their language ability. The fact that three individuals made this request for additional materials strongly indicates that participants were able to identify the unique value of the parallel corpus in helping them navigate meaning in Chinese through pinyin and their L1.
CHAPTER 6
CONCLUSION

As has been demonstrated, the parallel corpus tool provided a unique learning experience for learners, resulted in increased development of participants when applied as directed, and led to a positive learning experience for both learners and instructor. The following section first summarizes these findings, emphasizing the three primary research questions and their results. Next is a discussion of the implications of this research, looking at how the findings contribute to the body of research on parallel corpora as well as pedagogical practices. The following section then draws attention to the limitations of this study from a technological as well as a methodological perspective. After this, the final section examines potential directions for further research and application of parallel corpora, before concluding with a few final remarks.

6.1. Research Questions and Results

The principal objective in completing this research project was to develop and then implement a learning approach that would assist learners in more effectively and efficiently acquiring written Chinese. The parallel corpus functioned as a mediational tool that attempted to accomplish this task by making participant experience with learning Chinese less intimidating and their interaction with characters more organic. To be able to observe and analyze the overall efficacy of the parallel corpus and this approach to learning, three primary research questions were identified, which addressed

Learner Achievement, Pedagogical Application, and Learner Attitude/Perception:

- Learner Achievement – Will the Chinese/English parallel corpus aid students in learning Chinese?
- Pedagogical Implementation – How do the students and instructor implement the tool in the classroom throughout the semester?
Learner Attitude/Perception – How do students view the application and implementation of the tool?

To address these questions, both the product and process of the implementation of the parallel corpus were analyzed in order to observe how both instructor and participants adapted the tool, and then discern how these applications of the corpus led to (or not) development.

The results of the research demonstrated that integration of the parallel corpus led to substantial learner achievement and development when the participants adapted the parallel corpus as directed by the instructor. Isabel and Ellen provided the two most salient illustrations of this observation. While these two participants required very different levels of instructor mediation in learning to adapt the corpus, the end results of both individuals’ performance clearly demonstrated how the appropriate use of the tool led not just to acquisition of terms and constructs, but to functional understanding of the language. In Isabel’s case, this led to a contrasting performance during the final month of the study as she composed complex sentences expressing her own ideas, and ultimately functioning through the language to create her own meaning. While Ellen evidenced successful use of the tool throughout the study, the exponential progress she made by using the parallel corpus demonstrated not only its value, but also effectiveness for improving learner experience and development.

Other participants such as Gail, Leo, and Diane observed similar positive learning outcomes, but not quite as pronounced as Isabel and Ellen. While other participants’ applications of the parallel corpus did not clearly lead to improved development of their conceptualization of the Chinese language and its constructs, the consideration of their learning experiences demonstrated the importance of pedagogical implementation. The
way an individual chose to apply the parallel corpus greatly influenced their learning experience and level of development. Analysis of the learners who did not evidence observable improvement by using the tool revealed that they chose to use the parallel corpus differently than directed to simply search out answers and not to analyze the meaning and function of the language and specific constructs. For individuals such as Isabel, who began using the tool in this way, instructor mediation proved to remedy this issue and lead to marked improvement and development. For others, the level of instructor mediation did not affect their adaptation of the tool, as they still chose to use it differently.

Thus while instructor mediation and guidance in using the tool is essential, ultimately learners must also actively choose to use the tool as a mediational artifact in order for it to benefit their learning. Consideration of learner attitude/perception further substantiates this observation by revealing that even though several participants understood how the tool was meant to function they chose to use it differently. Hanna’s example clearly demonstrates this observation as she reported a positive learning experience with the parallel corpus and noted that while bilingual dictionaries “helped [her] find what new vocab items meant, [the] parallel corpus helped [her] figure out how to use them”. Although she verbalizes this correct understanding of the use of the corpus, her actual application of the tool was inconsistent with this observation – suggesting she understood its intended mediational function, but chose to use it differently. Despite this, and other similar observations where the corpus was not implemented as directed, the resounding learner attitude toward the parallel corpus was still very positive. The extent to which the corpus actually helped an individual improve did not affect assessments,
which indicated that users understood how the tool could benefit them in learning Chinese.

These observations suggest not only the actual value and potential of implementing parallel corpora in the classroom, but also the perceived value among learners. This attitude is also mirrored in the analysis of the instructor’s perception, as he notes the overall positive learning experience throughout the study and his motivation to use the parallel corpus to make challenging learning objectives more accessible to the participants. The observation of the instructor’s implementation of the parallel corpus also provides an example for teacher development and application that other instructors can refer to when considering the pedagogical adaptation of a parallel corpus or other language learning technology.

6.2. Implications for Pedagogy and Research

The results of this research carry several implications for both pedagogical and research applications of parallel corpora. From a pedagogical standpoint, the study provides a clear illustration of potential applications of parallel corpora. This serves to inform better practices of instructor adaptation of such technology, but also to highlight aspects of technology design that promote learner development. From a research perspective, this project answers the call made by several researchers to enrich the small body of research on parallel corpora in pedagogy, and provides a different theoretical approach than has been observed in prior parallel corpus studies (Barlow 2000; Johansson 2009; Wang 2001; Fan & Xu 2002).

The illustrative example of the implementation of the parallel corpus tool in this study provides researchers and teachers alike with an example to work from in
implementing similar technologies, and highlights many aspects of both learner and teacher experience that should be considered when using parallel corpora. Perhaps more importantly, though, are the implications for parallel corpus design for pedagogical applications. In reviewing the literature only two other studies were identified in which parallel corpora were created specifically for classroom use (Lavid, Hita, and Zamorano-Mansilla 2010; Bernardini 2003). Both of these studies demonstrated the benefit of using a parallel corpus designed for pedagogy as opposed to research, but focused primarily on improving interface design and simplifying user experience. Additionally, Bernardini (2003) adjusted texts to encourage greater levels of cultural understanding and awareness when using the English/Italian corpus. The current project considered these aspects, and went a step further by adapting additional language formats (pinyin and characters with tone marks) and other interactive features that were included specifically to benefit learners. While many of these adaptations were made to address unique challenges of Chinese language learners, the success of this approach does imply the potential for attempting similar adaptations in creating enhanced features for other language pairs that would better improve learner experience and capacity for language development.

These pedagogical implications, as well as the sociocultural theoretical approach taken in analyzing this study, serve to benefit and further develop the body of parallel corpus research. While reciprocal learning (Hunston 2002; Laviosa 2002), discovery learning (Bernardini 1996), and data driven learning (Frankenberg-Garcia 2005; Wang 2001) theories have been explored in studies adapting parallel corpora, this is the first study to the researcher’s knowledge that analyzes this language technology through the lens of SCT. Analysis through the SCT framework provides further insight into how
parallel corpora influence learner development as well as a theoretical basis for understanding the processes that lead to this development. The inherent make-up of parallel corpora present users with complete texts and concepts in the target language as well as their L1, and when effectively mediated individuals can come to develop conceptual knowledge of the target language through the established schemas of their L1. Better understanding this process of development can help guide improved mediational practices and corpus design to improve user experience and learner development.

6.3. Limitations

While the value of parallel corpora in pedagogy is evidenced through the current study, as well as others (Frankenberg-Garcia 2000; St. John 2001; Xu & Kawecki 2005), some limitations must be acknowledged. The Chinese/English parallel corpus was designed specifically for pedagogical use, but after studying its application there are a few design aspects of the technology that were potentially limiting to learner development and could be improved on. Additionally, limitations of the methodology and setting of the research must also be addressed.

From a technological perspective there were a few features that possibly limited the effectiveness of the parallel corpus. The participants identified some of these issues, while others became apparent through analysis. The lack of a drawing feature is perhaps the design issue that stands out most, as this one feature is what drew several participants (Brian, Frank, Adam, Hanna) away from the corpus to use other websites such as nciku.com. The drawing feature of other websites allows the user to draw a character they do not recognize in order to search for the term. Similar to the participants mentioned, Jack and Kate also showed a preference for using a drawing feature, but they
had an integrated keyboard input method in their iPads that allowed them to use the iPad character drawing function with the parallel corpus. This resulted in Jack and Kate evidencing a much stronger preference for the parallel corpus, and implies that the inclusion of this one feature would serve to better motivate use of the corpus tool, and improve user experience. Since the implementation of this feature, a similar keyboard input method has been added to Apple computer operating systems, and thus would have solved this issue for the setting in this study, but it would likely still be a feature worth adding for improved application.

Another design feature identified by Kate was the necessity to select the language format of input before being able to query a term in the corpus. For instance, pinyin would have to be explicitly selected before the user could search for a term using pinyin. Kate suggested that it would be more convenient and efficient if the parallel corpus automatically recognized the input method. This issue is quite minor, and could easily be changed, but it does serve to emphasize a greater limitation of the parallel corpus tool’s capacity to mediate learner development. The parallel corpus tool was purposefully designed to require participants to select a language format in order to search, so that correct tone mark indication could be required when the user was searching pinyin. This feature was viewed as beneficial in emphasizing mastery of tones and in mediating learner development to acquire this knowledge. This feature was observed fulfilling this intended purpose when Diane and Adam both used the feature to more efficiently search characters for terms they knew pinyin and tone for. This feature was beneficial for these two participants, but Kate (and possibly others) recognized it as a functional limitation.
While this issue seems quite minor, it does illustrate an overarching issue with developing a tool that better mediates learner development. The results of this study indicated that learners experienced improved development when they used the parallel corpus as directed. For individuals such as Diane, Leo, Gail, Isabel, and Ellen this led to positive outcomes. For participants that chose not to use the tool as directed, though, development could not be explicitly linked to adaptation of the corpus. Altering the parallel corpus to require participants to focus on learning through the Chinese (and not simply searching the English to find answers) could likely improve development among individuals such as the group of participants resistant to the parallel corpus. By removing the users ability to just search for answers using English, the parallel corpus could function more effectively as a mediational tool among a wider range of learners. Such alterations could be permanently adapted, or altered in the tool for specific activities in order to more effectively guide learners in their adaptation of the parallel corpus as a mediational tool.

One final limitation to address is that of the limited size and diversity of the participants. The limited number of participants does restrict the ability to derive statistical observations, and to generalize the findings. However, the qualitative analysis provides a detailed look at several individual learning experiences and does strongly imply the efficacy of this technology, while also highlighting several aspects of design and implementation that can be applied in future settings. Perhaps the biggest limitation, though, is that of diversity among participants. The high school students in this study were all tech savvy and very comfortable with tablets, computers, and several other digital learning tools before they were ever introduced to the parallel corpus. Other
populations of learners would likely witness a different learning curve in adapting the technology, and possibly report and experience diverse learning outcomes.

6.4. Future Directions

The future of parallel corpora research is expanding as an open field of experimentation and growth in both linguistics research and pedagogical application. As Johansson (2007) optimistically states, parallel corpora “have only been in use for some 10-15 years, but have already proved to be of great value… It is likely that we are only at the beginning of an exciting development. Much remains to be done” (p. 315). His optimism is not unique as many others (Wang 2002; Frankenberg-Garcia 2005; Xu & Kawecki 2005) share this enthusiastic outlook for the future of parallel corpus research. The potential applications of parallel corpus research extend far beyond the approaches that have been tackled in the literature, and the future of this field will undoubtedly reveal many innovative and practical methods to aid learners in the acquisition of language.

Findings from the current study suggest that future pedagogical applications of this technology should look beyond simply adapting extant corpora, and seek to further develop the technology to better suit language learner needs. The current study demonstrated how presenting Chinese language learners with text in additional formats (pinyin and characters with tone marks) led to improved outcomes and interactions with the parallel corpus. Similarly, Xu and Kawecki (2005) illustrated how using a Chinese/English/French trilingual corpus benefited their unique population of learners who were native Chinese speakers with advanced English proficiency learning French. What both cases illustrate is how looking beyond the standard design of parallel corpora, and incorporating additional languages, formats, or other features specific to a population
of learners can better serve that population. As each language is unique and presents learners with different obstacles, the design of each parallel corpus could also address these unique features.

Beyond addressing unique learning obstacles in various language pairings, a clear direction for further development of parallel corpora is in adapting multimedia content. Several participants in the current study suggested that the parallel corpus could be improved by adding audio features and video clips. Adding these and other multimedia features such as images and animations would likely improve learner experience by making the learning process more multimodal. From the perspective of the current study, the purpose in integrating the parallel corpus was to improve learners’ acquisition of written Chinese and development of conceptual knowledge through the language. Providing learners with the ability to draw on parallel schemas from their L1 allowed them to more readily conceptualize content through in the target language. Going one step further to integrate audio, video, imagery, and any other multimedia form that aided this process of development would likely lead to improved outcomes and is a viable path for future exploration.

Finally, additional resources have become available since the parallel corpus was developed for the current study. Specifically, linguee.com, a website that allows users to search parallel concordances in multiple language pairings, has developed substantially and is now available in Chinese/English. The corpus designed in this study was limited by size and scope, and a clear next direction is in expanding the content available to users. While linguee.com provides far more content, its design is still limited in its capacity to mediate learner experience. The researcher is currently working to explore
pedagogical applications of the linguee resource, and working together with the web-based company to explore future development of this resource to make it more functional as learning resource. The intent is to both enhance the current tool, and then be able to study a much broader population of users in how they use parallel corpus technology.
REFERENCES


language acquisition at the beginning level. *Dimension, 1*, p. 31-48.


work and new directions. In S. Botley, T. McEnery, & A. Wilson (eds.).


Corpus-Based Approaches to English Language Teaching. New York: Continuum International Publishing Group, p. 138-148.


(2007c). *Common Knowledge About Chinese History* (中国历史常识). Beijing, PRC:
Gāoděng Jiàoyù Chūbānshè (高等教育出版社).


APPENDIX

Learning Log Prompt

- What resources did you use in completing this assignment (book, dictionary, technology, etc.)?

- What resource did you find most helpful? Why?

- What was most challenging about this assignment? Why?

- What resources might make this less challenging in the future?

- What did you find most beneficial to you on completing this task?

- Please list any other comments you might have about your learning experience with this assignment.
End of Semester Questionnaire

- What did you find most enjoyable about the class during the semester?

- What did you find most challenging about the class during the semester?

- What was your favorite lesson or topic that we covered?

- What resources did you find most helpful to you in your language study? Why?

- What would you change about the class to make it better?

- What resources did you find most helpful to you in learning Chinese throughout the semester? Why?

- Of the resources that you listed above, how did you use them? Please describe.

- If you could have your ‘ideal’ language-learning tool or resource for learning Chinese – what would it be like? Please describe it in detail.

- Please rate your response to the following questions on a scale of 1-7, with one being the lowest and seven being the highest
1 – How useful was the parallel corpus tool for you throughout the semester?

1 (Not useful)  2  3  4  5  6  7 (Very Useful)

2 – How often did you use the parallel corpus tool in learning and studying Chinese?

1 (Never)  2  3  4  5  6  7 (Very Frequently)

3 – How likely are you to use the parallel corpus tool in the future?

1 (Not likely)  2  3  4  5  6  7 (Very Likely)

Please answer the following questions.

4 – What did you like about the parallel corpus tool?

5 – What didn’t you like about the parallel corpus tool?

6 – How did you use the parallel corpus tool? Please describe in detail and give examples.

7 – How could the parallel corpus tool be improved?

8 – Please list any further comments or suggestions you have about the parallel corpus tool.
Autoethnography

Purpose: The auto ethnography will include entries on how the instructor prepares for each class and prepares to use the tool in class, as well as post-instructional accounts that detail how the class (and daily implementation of the tool) went.

1 – Did you plan/prepare to use the tool in class today?
   □ Yes □ No

2 – Did you use the tool in class?
   □ Yes □ No

3 – If it was used, how was it accessed?
   □ IPads □ Computers/Computer Lab □ Other

4 – How did you prepare to use the tool for class? What activities were planned? How did you plan on using the tool to aid in these activities?

5 – Did any modifications/adjustments need to be made to the tool before using it in class?

6 – How was it actually used in class? Was it implemented as planned, or used in a different manner?

7 – What was the learner experience like? Did it make the activity more effective, challenging, something else?

8 – What challenges did you and/or the learners encounter, if any?

9 – What changes/adjustments were made to the tool following the class?

10 – Any other thoughts, feelings, impressions.
### Hadjerrouit’s (2010) WBLR Framework

#### Table 1: Conceptual framework for identifying the key points of WBLRs

<table>
<thead>
<tr>
<th>FACTORS &amp; KEY POINTS</th>
<th>ISSUES AND QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WBLR features</strong></td>
<td></td>
</tr>
<tr>
<td>Pedagogy</td>
<td>Does the school/class have pedagogical and learning strategies?</td>
</tr>
<tr>
<td>Technology</td>
<td>Is the existing software infrastructure able to support WBLRs?</td>
</tr>
<tr>
<td>Content</td>
<td>Which content of the subject matter is relevant?</td>
</tr>
<tr>
<td><strong>Usability criteria</strong></td>
<td></td>
</tr>
<tr>
<td>Technical usability</td>
<td></td>
</tr>
<tr>
<td>Content design</td>
<td>Is the content easy to read and to access?</td>
</tr>
<tr>
<td>Page design</td>
<td>Are the pages well-structured?</td>
</tr>
<tr>
<td>Site design</td>
<td>Is the navigation intuitive and easy to follow?</td>
</tr>
<tr>
<td>Pedagogical usability</td>
<td></td>
</tr>
<tr>
<td>Understandability</td>
<td>Is the content easy to understand?</td>
</tr>
<tr>
<td>Added value</td>
<td>Does the WBLR bring something better than existed already?</td>
</tr>
<tr>
<td>Goal-orientation</td>
<td>Are the learning goals set by the teacher and curriculum?</td>
</tr>
<tr>
<td>Time</td>
<td>Is the time needed for learning taken into consideration?</td>
</tr>
<tr>
<td>Interactivity</td>
<td>How good is the WBLR's feedback on the students’ actions?</td>
</tr>
<tr>
<td>Multimedia</td>
<td>Does the WBLR combine a number of multimedia elements?</td>
</tr>
<tr>
<td>Motivation</td>
<td>Is the WBLR adapted to the students’ characteristics?</td>
</tr>
<tr>
<td>Differentiation</td>
<td>Can the WBLR be tailored to the student?</td>
</tr>
<tr>
<td>Flexibility</td>
<td>Does the WBLR appear open and flexible?</td>
</tr>
<tr>
<td>Autonomy</td>
<td>Does the WBLR support autonomy of learning?</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Is the WBLR designed to support collaboration?</td>
</tr>
<tr>
<td>Variation</td>
<td>Is the WBLR designed to support variation?</td>
</tr>
<tr>
<td><strong>Context of use/evaluation</strong></td>
<td></td>
</tr>
<tr>
<td>Non material milieu</td>
<td></td>
</tr>
<tr>
<td>Textbooks</td>
<td>Is it intended to use textbooks in combination with the WBLR?</td>
</tr>
<tr>
<td>IT infrastructure</td>
<td>Does the IT infrastructure provide sufficient support?</td>
</tr>
<tr>
<td>Curriculum</td>
<td>Does the WBLR fit the curriculum?</td>
</tr>
<tr>
<td>Subject matter</td>
<td>How is the subject implemented through the WBLR?</td>
</tr>
<tr>
<td>Material milieu</td>
<td></td>
</tr>
<tr>
<td>Fellow learners</td>
<td>Is it intended to use the WBLR collaboratively?</td>
</tr>
<tr>
<td>School teachers</td>
<td>What is the school teacher’s attitude to the WBLR?</td>
</tr>
<tr>
<td>Trainee teachers</td>
<td>What is the role of trainee teachers in the use and evaluation?</td>
</tr>
<tr>
<td>School leaders</td>
<td>What is the school policy regarding WBLRs?</td>
</tr>
</tbody>
</table>

(Hadjerrouit 2010, p.61)
Brody Bluemel has been granted full permissions to include complete portions of text from the article he authored entitled, “Learning in Parallel: Using parallel corpora to enhance written language acquisition at the beginning level”. This article was published in volume 1 of the 2014 edition of *Dimension*.

David Jahner  
Executive Director, Southern Conference on Language Teaching  
P.O. Box 33615  
Decatur, GA 30033  
Email: scoltdj@gmail.com
Brody T. Bluemel

Departments of Applied Linguistics and Asian Studies
The Pennsylvania State University
304 Sparks Building
University Park, PA 16802
brody.bluemel@gmail.com

Education

  Dissertation: Parallel Corpora and Pedagogy: Enhancing Chinese Foreign Language Learning Experience through Parallel Corpus Technology
  Committee: Xiaofei Lu (advisor), Ning Yu, Karen Johnson, and Robert Schrauf.
♦ B.A., English & Language Acquisition, Drake University, 2008.

Publications and Technology Development


Certifications

♦ Teaching with Technology Certificate. The Pennsylvania State University. (December 2013).
♦ SUV.3 Phonetik Schein (Certification for teaching phonetics and oral language). Received from Johannes-Gutenberg University, Mainz, Germany (July 2010).

Honors, Awards, Grants, and Fellowships

♦ Confucius Institute Grant, The Pennsylvania State University, 2013
♦ Gil Watz Fellowship, The Pennsylvania State University, 2012-2013
♦ Karako Fellowship and Grant, Summer 2012
♦ Confucius Institute Grant, The Pennsylvania State University, Spring 2012
♦ ChinaSprout Presentation Prize, The Annual Meeting of the Chinese Language Teachers Association (CLTA-13), 2013
♦ Research Travel Grant, Third International Summer School for China Studies, Summer 2012

Academic Positions

2010-2015  Teaching Assistant in Asian Studies, Applied Linguistics, German Department and ESL Program - The Pennsylvania State University
2012-2014  High School Chinese Language Instructor and Program Advisor – State College Area High School
2011-2012  Interim Director of ESL Program – The Pennsylvania State University